ALPHATECH, INC.

AD-A259 423

TR-565

INCIPIENT FAILURE DETECTION USING WAVELETS

December 1992

Prepared by:

Robert R. Tenney
Alonso E. Rhenals
James C. Deckert
Alan S. Willsky

Submitted to:

Dr. Kam Ng
Office of Naval Research
Code 122, Room 528
800 N. Quincy St., BCT #1
Arlington, VA 22217-5000

Contract No. N00014-92-C-0040
CDRL A002

ALPHATECH, Inc.
50 Mall Road
Burlington, MA 01803-4562
(617) 273-3388

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
Dec 92

3. REPORT TYPE AND DATES COVERED
Final Report (20 Apr 92 - 31 Oct 92)

4. TITLE AND SUBTITLE
Incipient Failure Detection Using Wavelets

5. FUNDING NUMBERS
N00014-92-C-0040

6. AUTHOR(S)
Robert R. Tenney
James C. Deckert
Alonso E. Rhenals
Alan S. Willsky

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
ALPHATECH, Inc.
Executive Place III
50 Mall Road
Burlington, MA 01803-4562

8. PERFORMING ORGANIZATION REPORT NUMBER
TR-565

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; SBIR report, distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This project demonstrated the feasibility of incipient fault detection for vibrating systems not only for bench test conditions (helicopter gearbox) but also for mild operating conditions (condensate and fire pumps). We employed only one of the several available data channels from the multiple accelerometers mounted on the case of each equipment type. We used a continuous wavelet transform (CWT), based on the Kiang wavelet, to select features for an artificial neural network (ANN) classifier. The CWT provided enough visibility into the fault signals to allow us to reduce the size of the feature set to 10 – 15 features. We used a low-dimensional, conventional ANN classifier with rejection of ambiguous classifications. We achieved 0.000 probability of false alarm, 0.000 probability of missed detection, and < 0.04 probability of deferral (to a subsequent feature vector) for all three data sets. The major product of this work is a single methodology to identify robust features that lead to these performance levels.

14. SUBJECT TERMS
Failure Detection, Wavelet Transforms, Neural Networks, Failure Classification, Signal Processing, Vibration Analysis

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
UNCLASSIFIED

18. SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

19. SECURITY CLASSIFICATION OF ABSTRACT
UNCLASSIFIED

20. LIMITATION OF ABSTRACT

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

CONTENTS

Section		Page
LIST OF FIGURES		iii
LIST OF TABLES		iii
COLOR PLATES		iv
1.	INTRODUCTION	1
1.1	Identification and Significance of the Problem	1
1.2.	Objectives of Phase I	2
1.3	Overview of Phase I Results	4
1.4	Report Organization	4
2.	PHASE I TECHNICAL EFFORT	5
2.1	Data Used in Phase I	5
2.2	System Structure	7
2.3	Wavelet-Based Tunable Preprocessor	9
2.3.1	The Continuous Wavelet Transform	10
2.3.2	Smoothing the CWT	11
2.3.3	Changing the Wavelet Basis	12
2.3.4	Channel Selection	12
2.3.5	Gearbox Fault Signatures	13
2.3.6	Gearbox Fault Masks	13
2.3.7	Condensate Pump Signatures	13
2.4	Feature Separation	13
2.4.1	Gearbox Separation	15
2.4.2	Condensate Pump Separation	16
2.4.3	Fire Pump Separation	16
2.4.4	Artifacts in Training Data	17
2.4.5	Guidelines for Finding Robust Feature Sets	17
2.5	Artificial Neural Network Classifier	18
2.6	Fault Detection and Identification Results from Phase I	20
3.	CONCLUSIONS AND RECOMMENDATIONS	22
3.1	Conclusions from the Phase I Effort	22
3.2	Phase II Recommendations	23
REFERENCES		26
ADDITIONAL RELEVANT PUBLICATIONS BY PHASE I TEAM MEMBERS		27
SELECTED BIBLIOGRAPHY ON FAULT DETECTION/ISOLATION AND WAVELETS		30
APPENDIX A: THE CONTINUOUS WAVELET TRANSFORM		33

LIST OF FIGURES

Number		Page
2-1	Incipient Fault Detection and Classification System Structure	7
2-2	Tunable Feature Extractor	9
A-1	The Haar and Kiang Wavelets	33
A-2	Dilated Translates of the Haar Wavelet	34
A-3	Typical Subdivision by Wavelets of the Time-Frequency Plane	34
A-4	Classification of Wavelets	35

LIST OF TABLES

Number		Page
2-1	CONDENSATE PUMP INFORMATION	6
2-2	FIRE PUMP INFORMATION	6
2-3	PHASE I CLUSTER SEPARATIONS, HELICOPTER DATA	14
2-4	NUMBER OF PROCESSING ELEMENTS PER ANN LAYER	18
2-5	PREPROCESSOR TIME CONSTANTS AND FEATURE VECTOR RATES	20
2-6	PHASE I PERFORMANCE RESULTS	21

COLOR PLATES

Plate A. Effect of Smoothing the Continuous Wavelet Transform

Plate B. Effect of Decreasing the Kiang Wavelet Frequency Resolution

Plate C. Condensate Pump Axial and Radial Data Channels in the CWT Domain

Plate D. Helicopter Gearbox Fault Condition Signatures in the CWT Domain

Plate E. Helicopter Gearbox Signatures After Masking Out Low-Level Energy

Plate F. Condensate Pumps Signatures in the CWT Domain

Plate G. Helicopter Gearbox Feature Cluster Separation

Plate H. Condensate Pump Feature Cluster Separation

Plate I. Fire Pump Feature Cluster Separation

Plate J. Artifacts in Training Data—Helicopter Gearbox

Plate K. Continuous Wavelet Transforms of a Pulse and a Sine

Plate L. Continuous Wavelet Transforms of Pulse Plus Sine and Two Sines

Plate M. Continuous Wavelet Transforms of Poisson and Gaussian White Noises

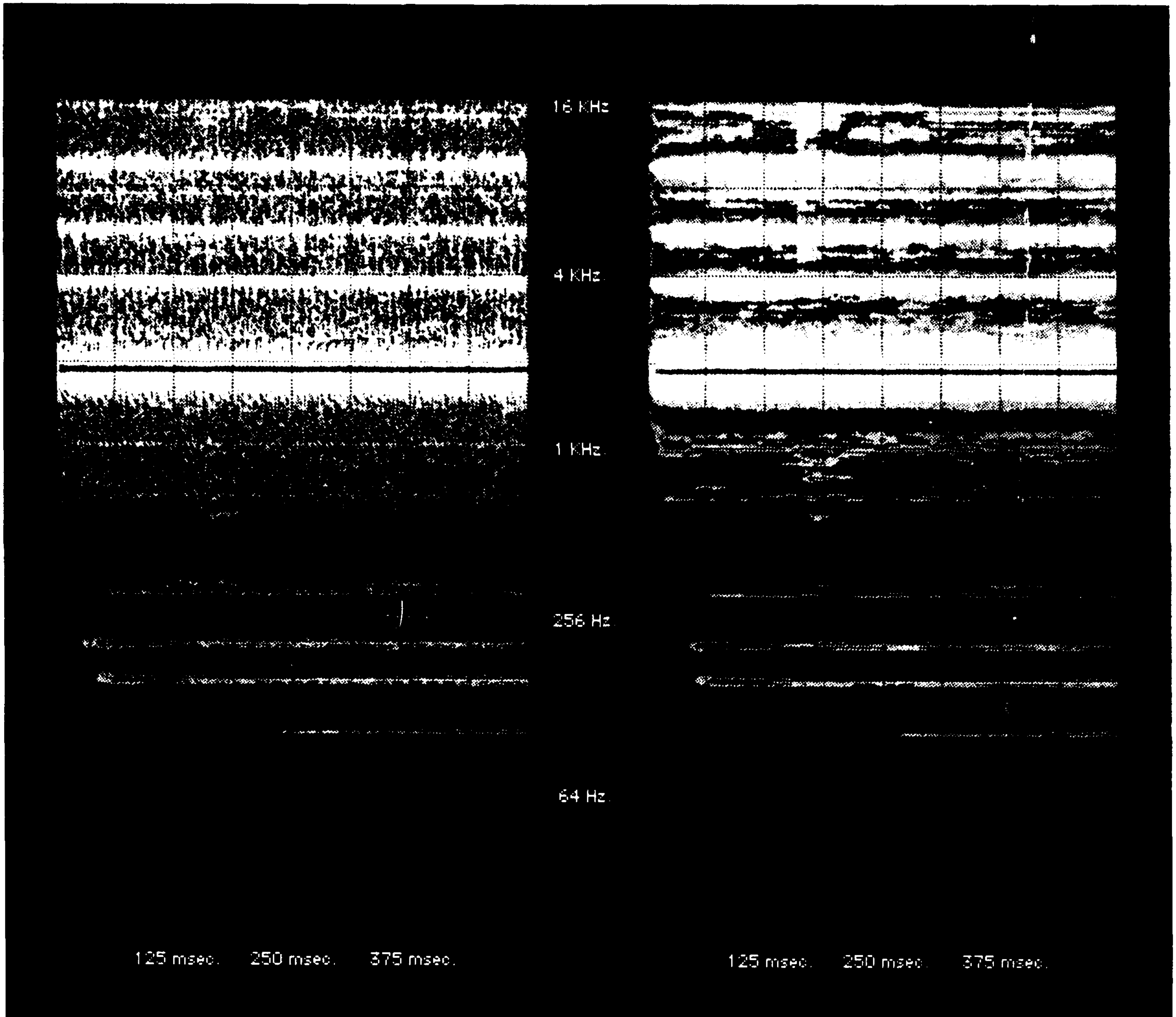

Plate A. Effect of Smoothing the Continuous Wavelet Transform

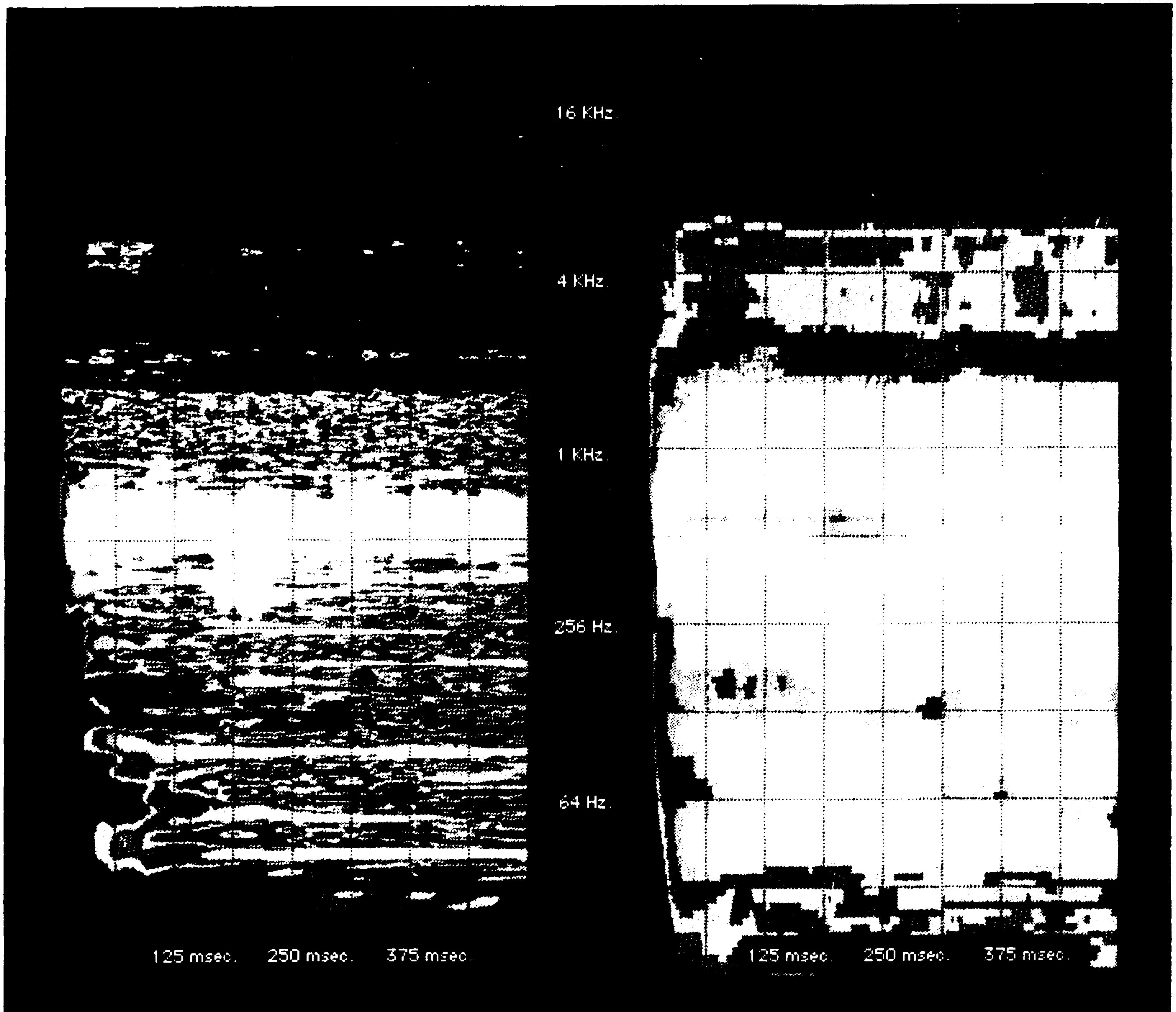

Plate B Effect of Decreasing the Kiang Wavelet Frequency Resolution

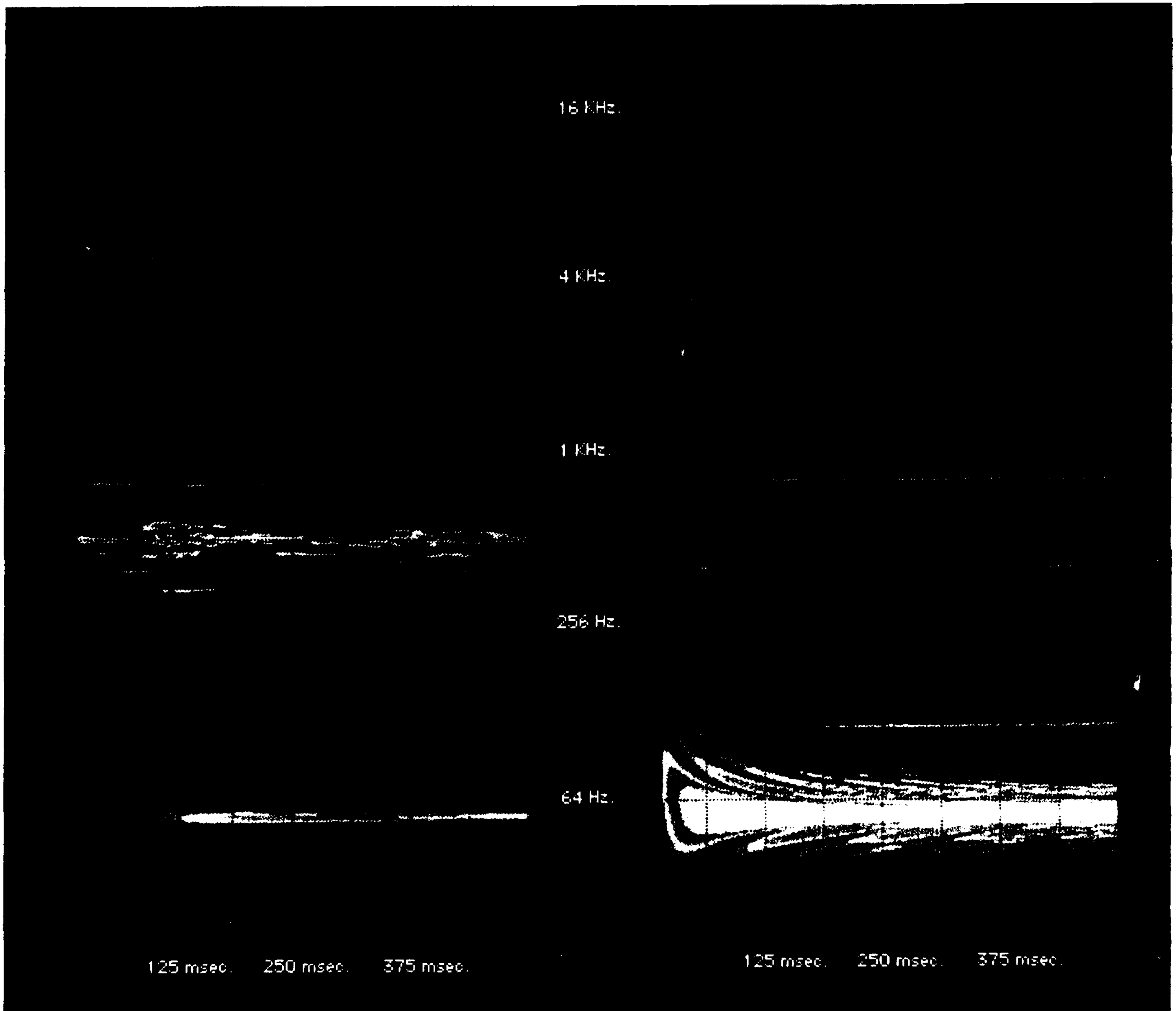

Plate C. Condensate Pump Axial and Radial Data Channels in the CWT Domain

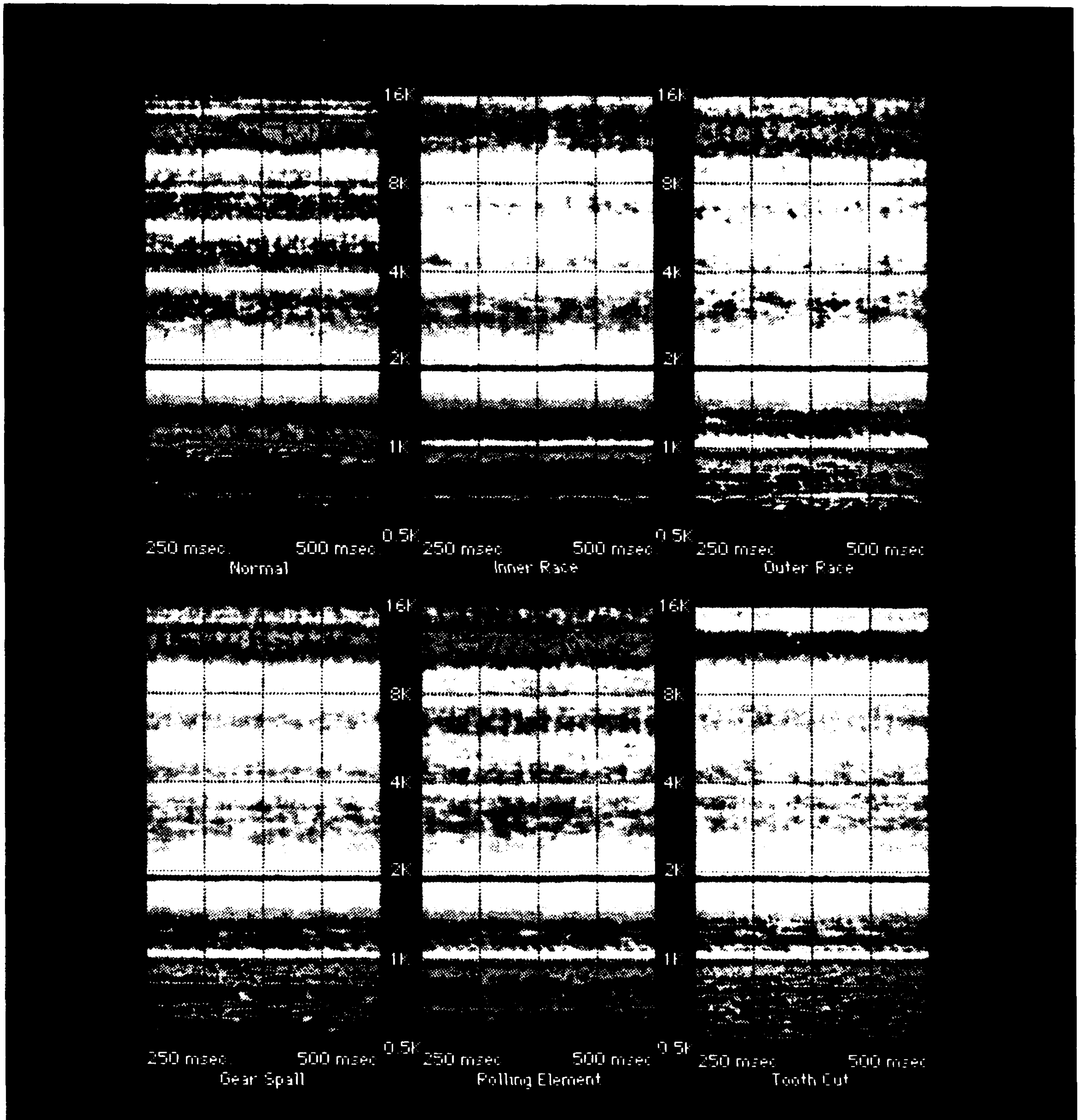

Plate D. Helicopter Gearbox Fault Condition Signatures in the CWT Domain

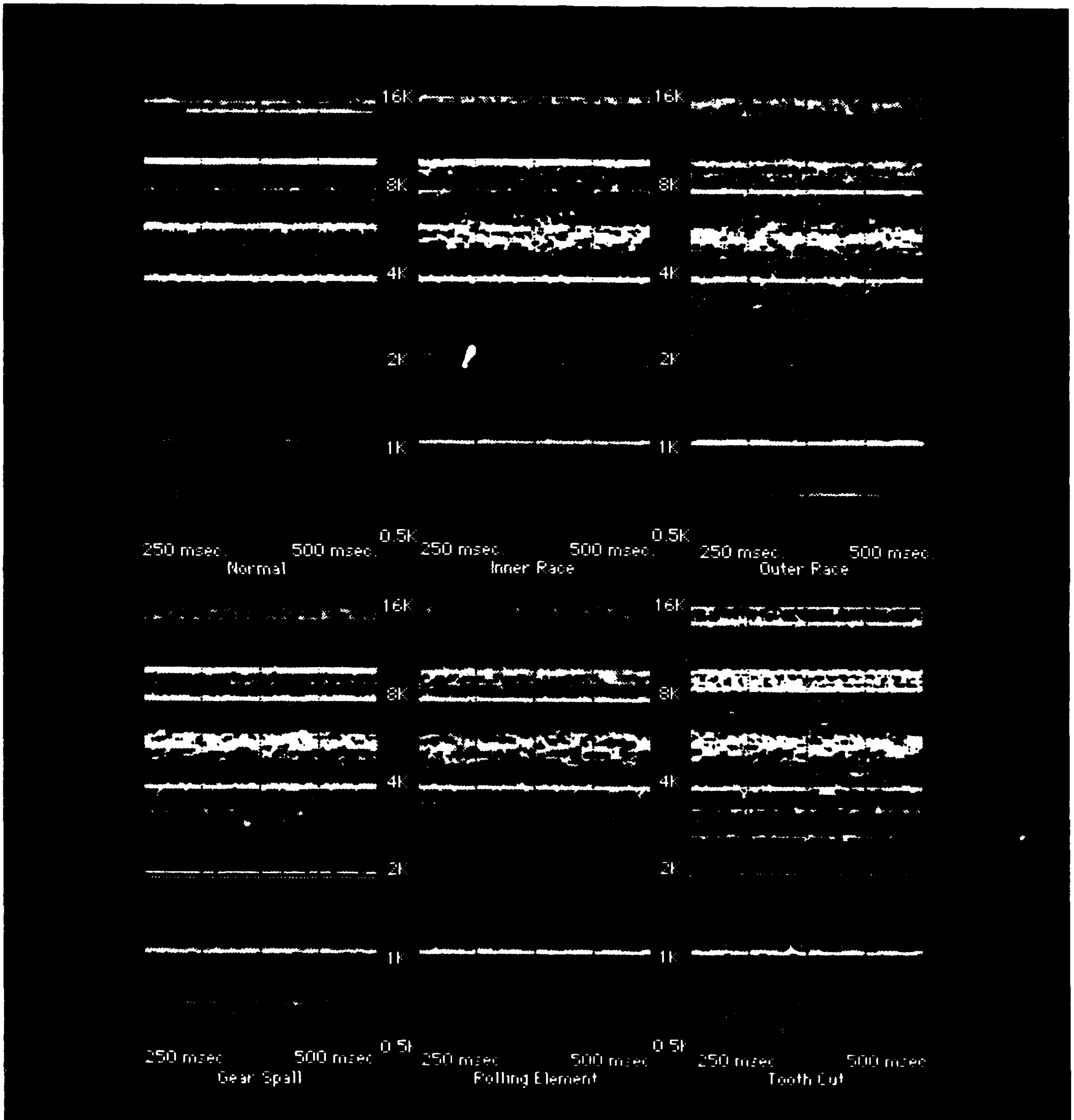

Plate E. Helicopter Gearbox Signatures After Masking Out Low-Level Energy

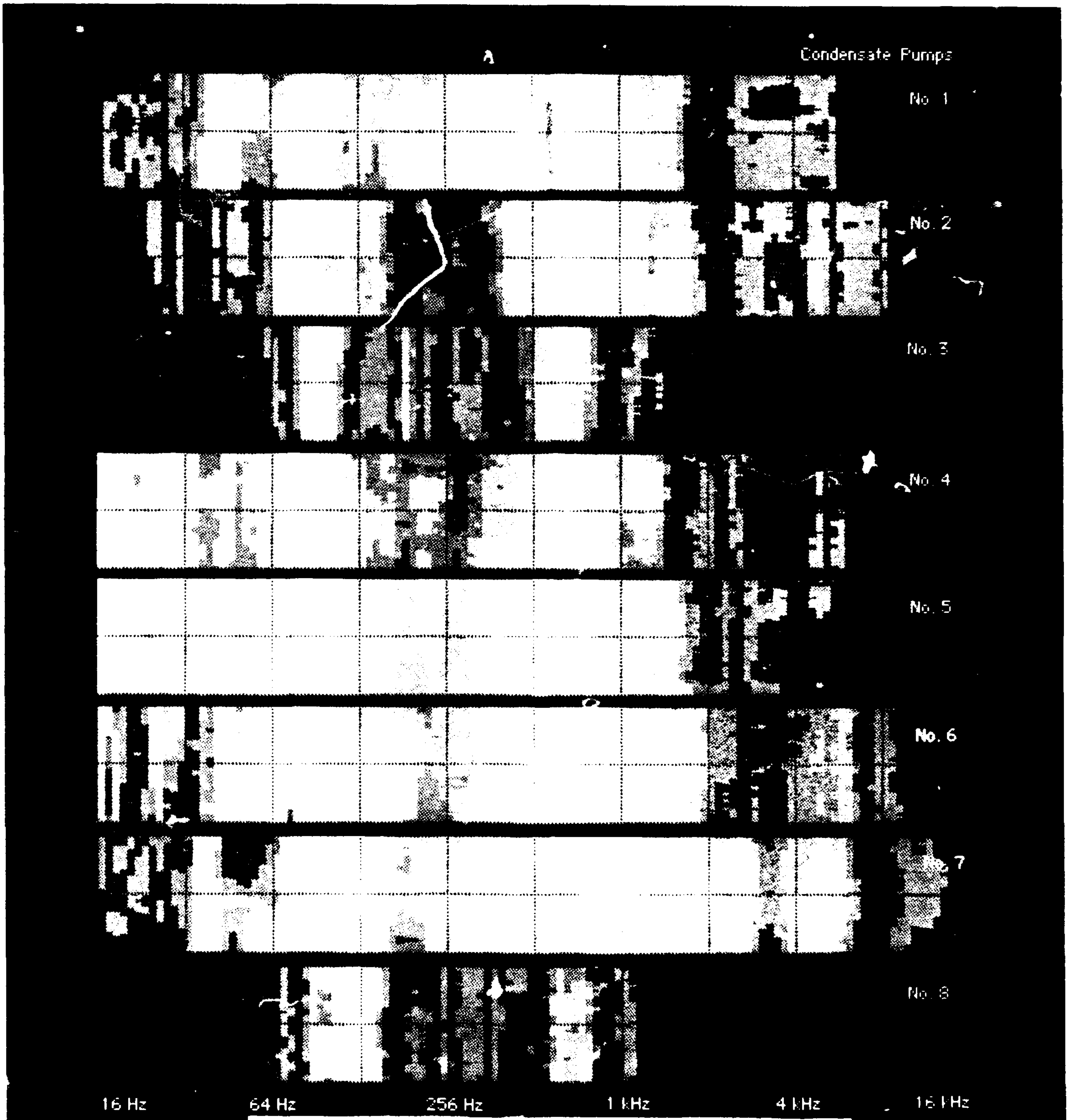

Plate F. Condensate Pumps Signatures in the CWT Domain

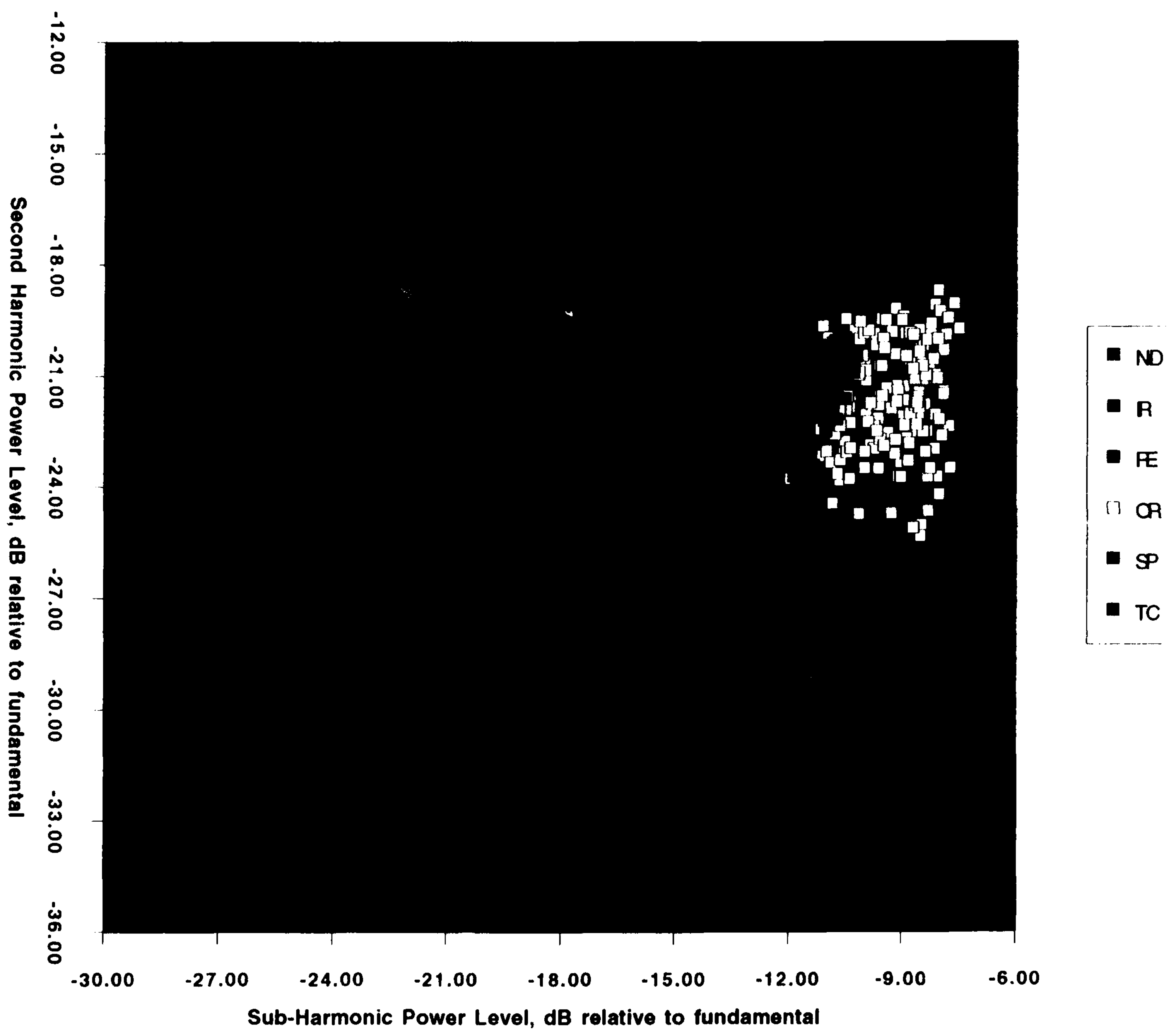

Plate G. Helicopter Gearbox Feature Cluster Separation

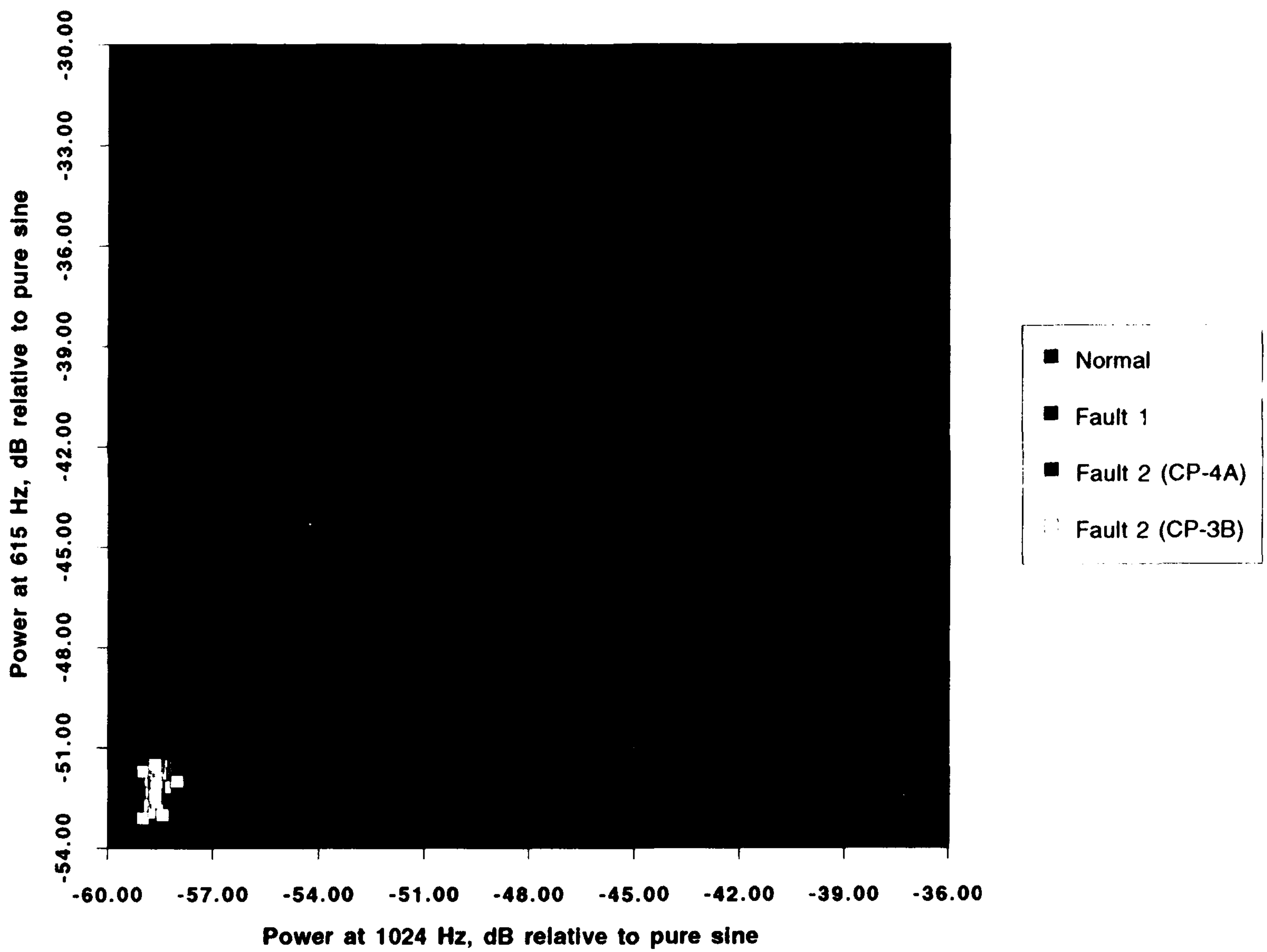

Plate H. Condensate Pump Feature Cluster Separation

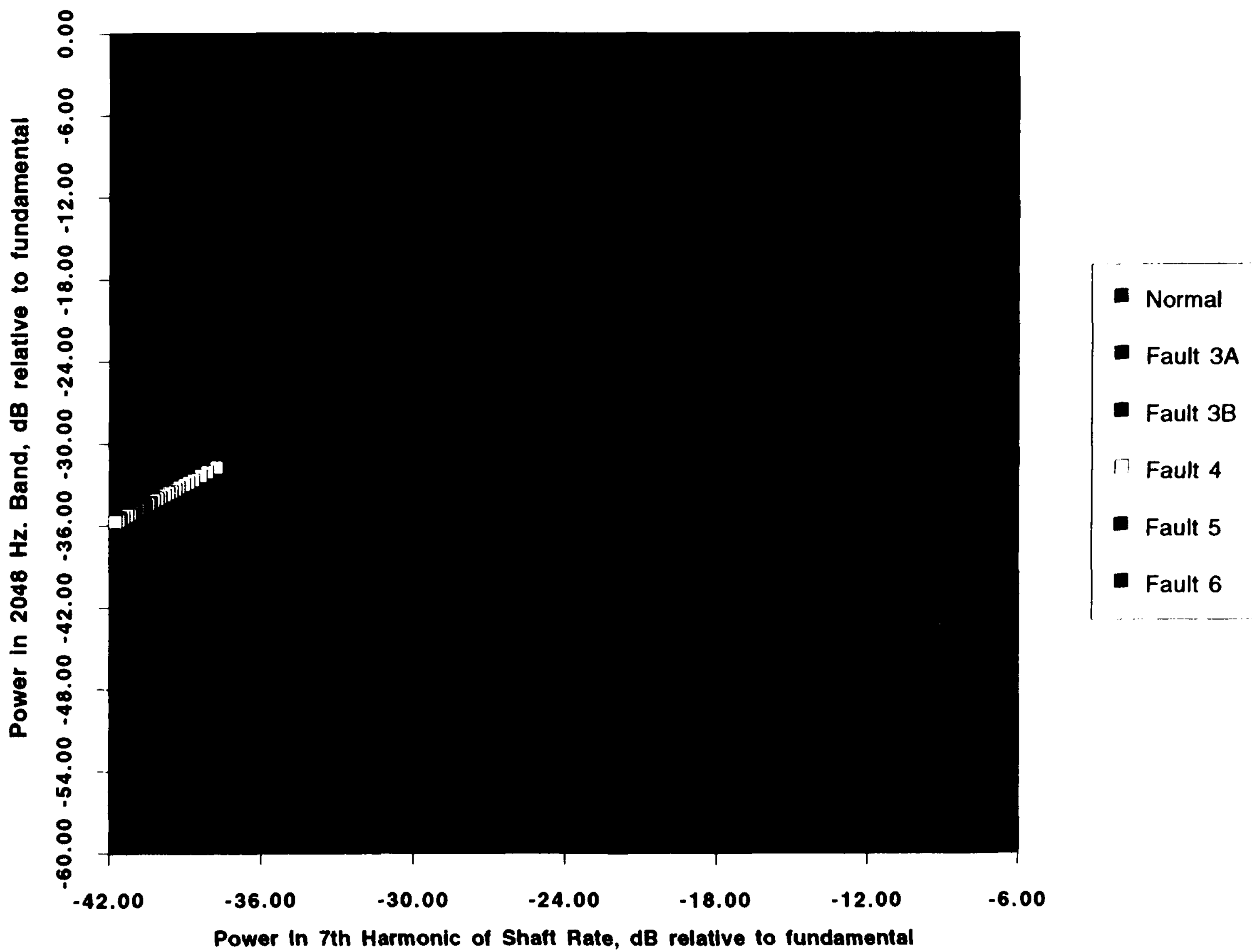

Plate I. Fire Pump Feature Cluster Separation

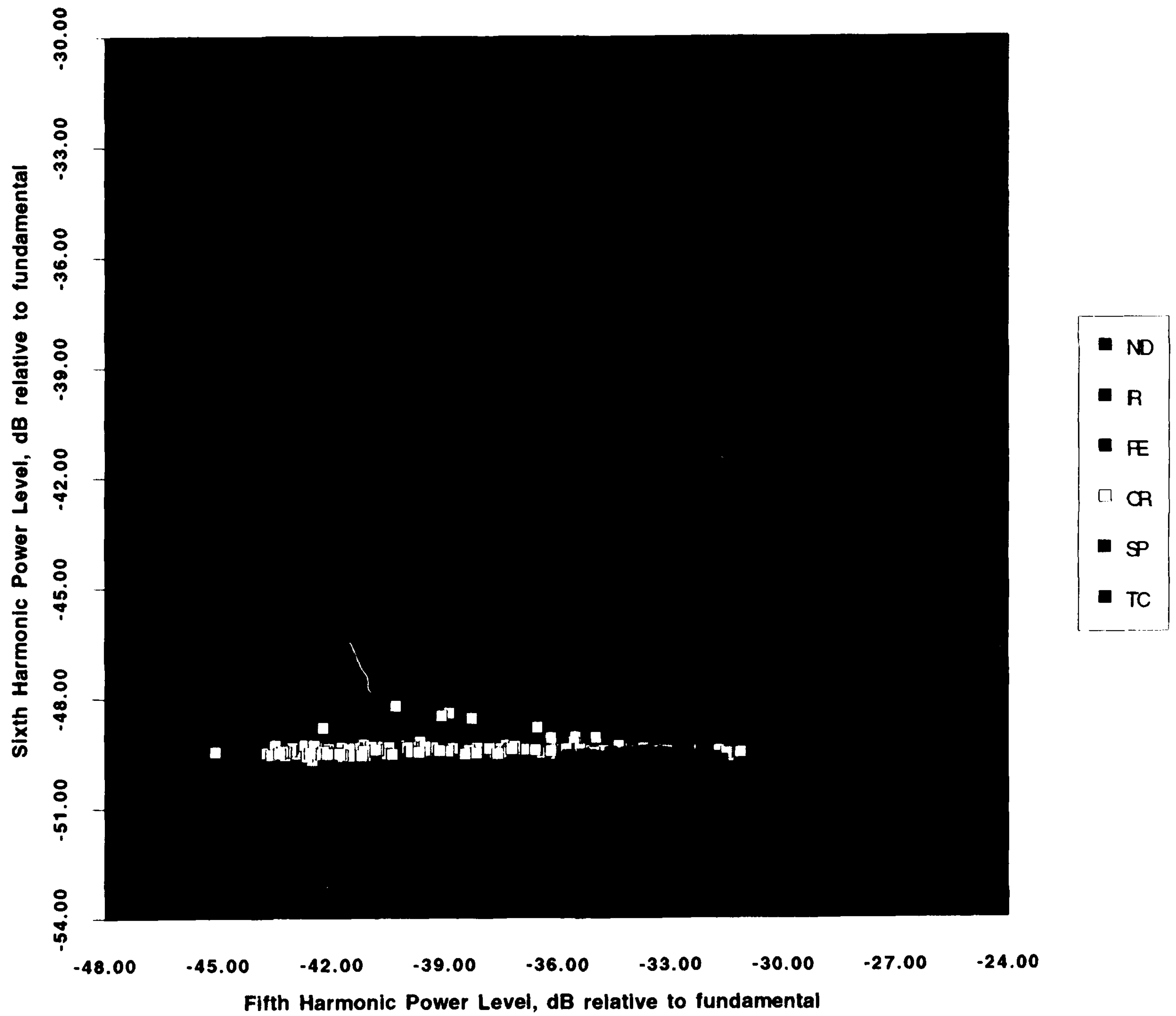

Plate J. Artifacts in Training Data—Helicopter Gearbox

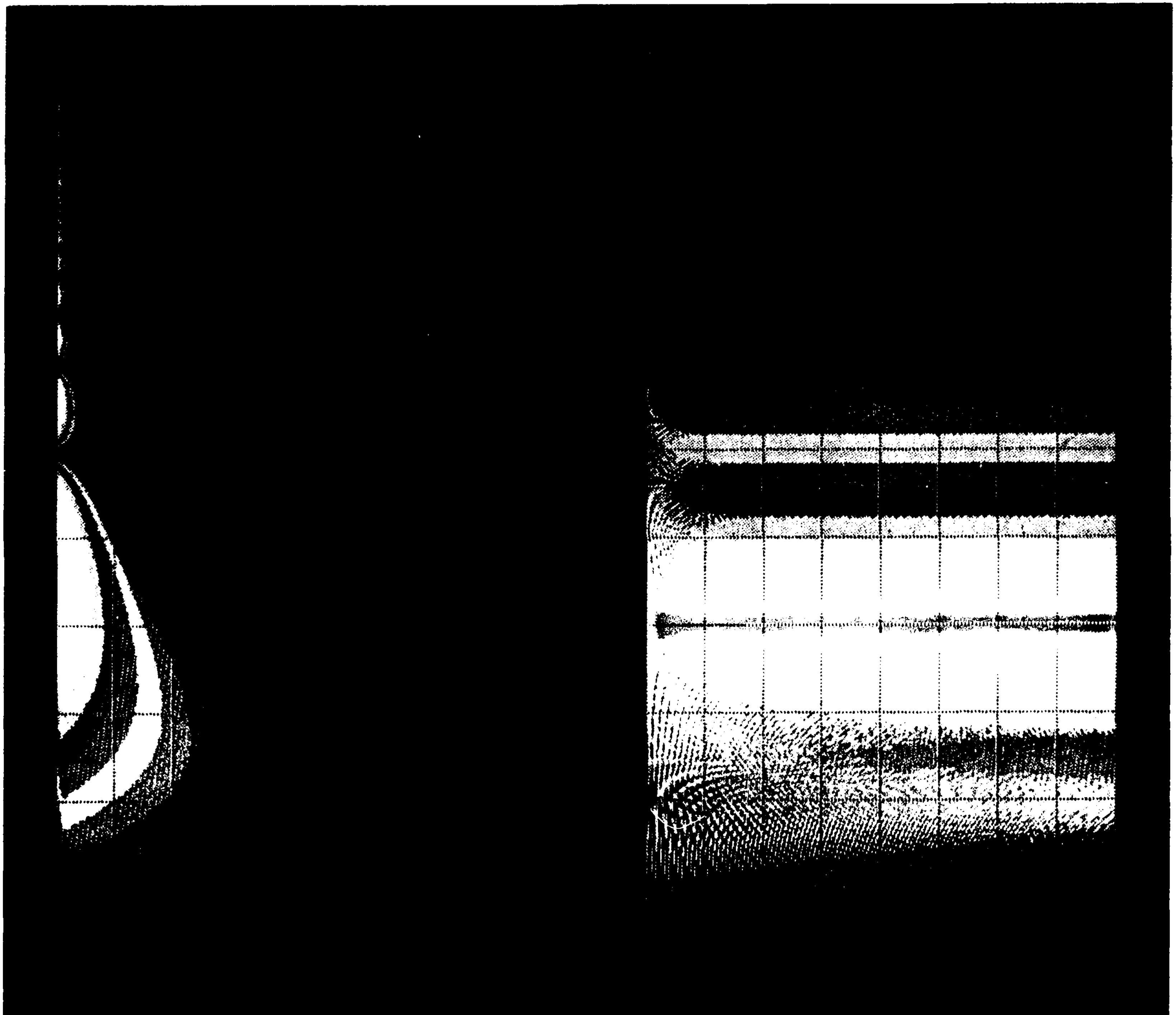

Plate K Continuous Wavelet Transforms of a Pulse and a Sine

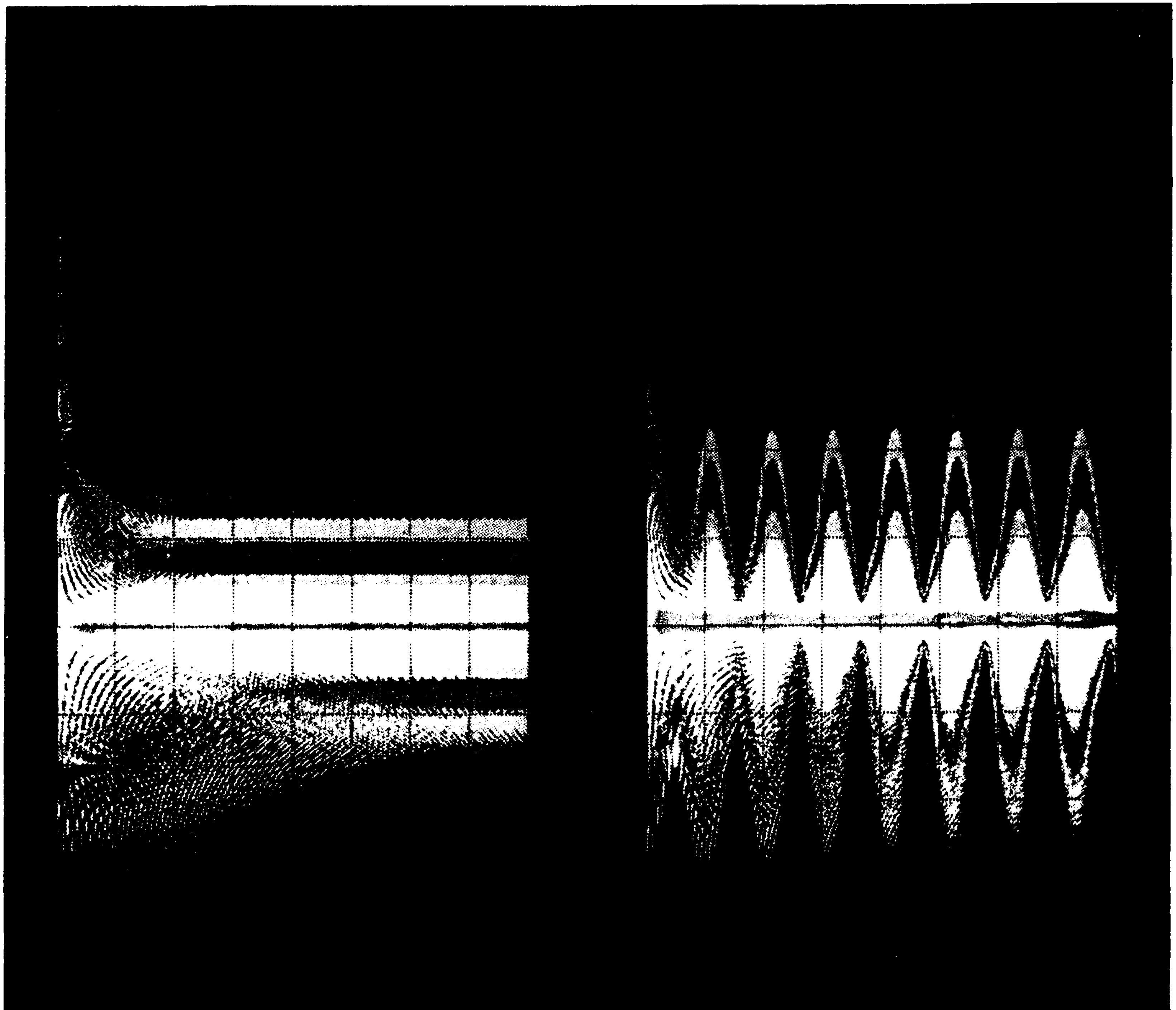

Plate L. Continuous Wavelet Transforms of Pulse Plus Sine and Two Sines

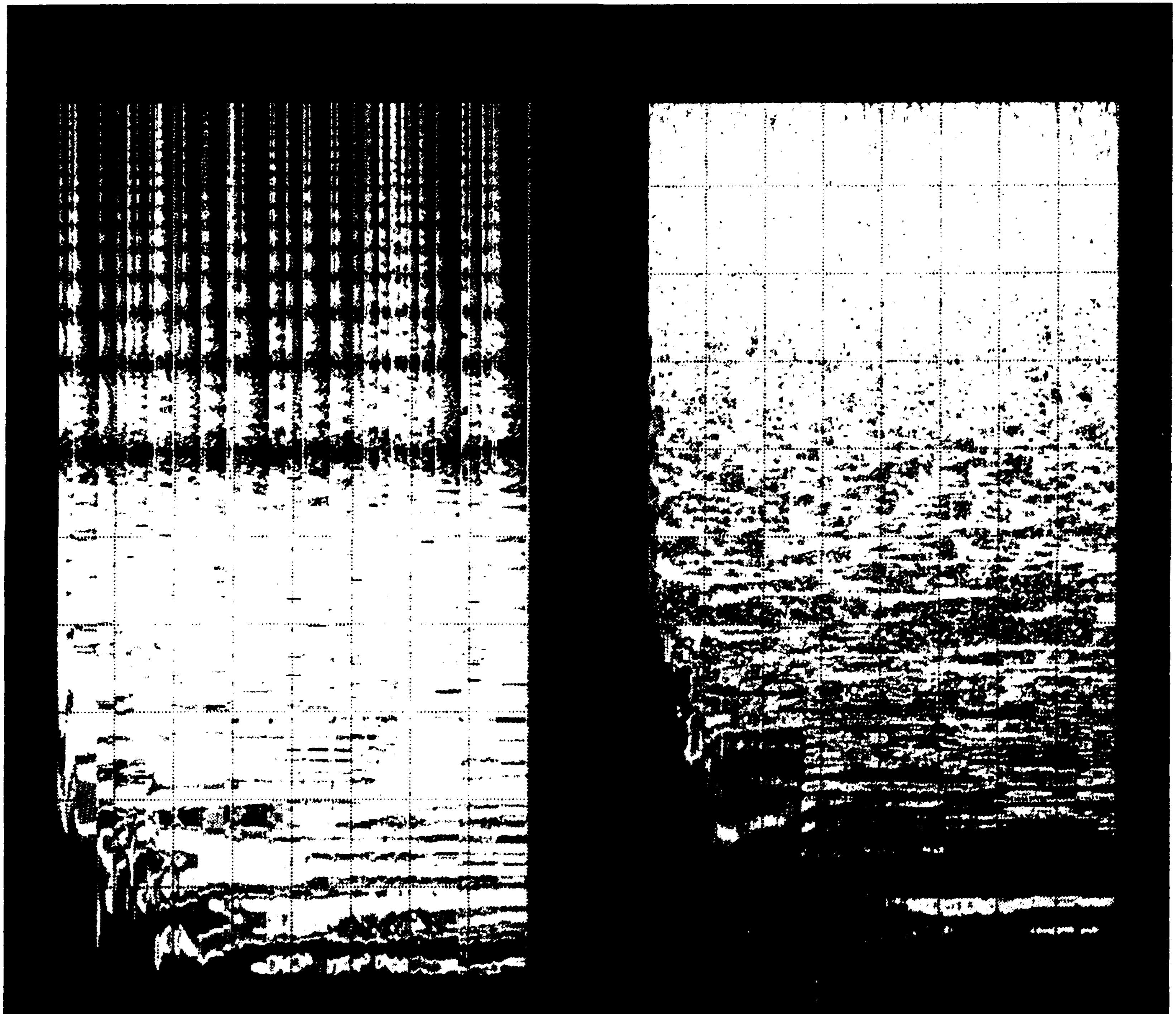

Plate M. Continuous Wavelet Transforms of Poisson and Gaussian White Noises

SECTION 1

INTRODUCTION

This report presents the results of a Phase I Small Business Innovation Research project funded by the Office of Naval Research in topic N91-297, Conditioned-Based Machinery Maintenance. Details of our proposed follow-on work for Phase II are given in ALPHATECH (1992).

1.1 IDENTIFICATION AND SIGNIFICANCE OF THE PROBLEM

The timely and reliable detection of changes in the dynamic behavior of complex systems and signals is a problem of considerable importance in a vast array of military and civilian applications. As we continue to place increasingly demanding objectives on system performance, cost, and reliability, the needs for and requirements on such detection methods grow commensurately. For example, the increasing role of and reliance on computer control—for the fly-by-wire control of advanced high-performance aircraft and helicopters, the navigation of autonomous vehicles, etc.—makes the detection of system anomalies essential, since by their very nature such automatic systems simply do not have the luxury of relying on the extraordinary but workload-limited detection capabilities of their human pilots. Also, the cost of modern-day military systems are such that there are tremendous payoffs to be gained if the availability of a weapons system is improved, or its life cycle cost reduced.

These objectives provided much of the motivation for the development of self-repairing flight control system concepts (Weiss and Hsu, 1987) for the in-flight detection of battle damage and sensor and actuator failures in advanced aircraft in order: 1) to facilitate control system reconfiguration to allow mission completion (or at least the safe return of the vehicle), and 2) to provide early diagnosis of problems that could then speed up the maintenance process and reduce turn-around time.

Furthermore, the reliable detection of component damage or failure can have a dramatic effect on the cost of maintaining and/or replacing an advanced military vehicle such as a helicopter, ship, or fighter. Specifically, the total cost of such a system is so high that the objective of avoiding system loss due to an undetected failure in some component places severe demands on the overall reliability of components and their monitoring systems.

Moreover, this need for reliability has typically led to the adoption of extremely conservative maintenance and replacement procedures: components are automatically replaced after time in service reaches a prescribed limit, usually taken to be significantly less than their expected failure times. Thus the availability of advanced and reliable fault detection systems offers the promise not only of improved system reliability but also the possibility of increasing component time in service by detecting the onset of problems and thus allowing "retirement for cause" rather than the more expensive present practice of replacing components whether they need it or not.

These and a variety of other factors and applications have led to considerable research and development activity over the past 20 years resulting in an array of detection and diagnosis methods (see, for example, the widely referenced surveys Willsky (1976) and Basseville (1987)) providing us with an analytically sound, proven-in-practice foundation from which to pursue the new challenges arising as we push harder on the envelope of performance, reliability, and cost. Moreover, in the past few years significant new methods of signal analysis and pattern recognition (in particular, wavelet transforms and artificial neural networks) have been developed offering the promise of adding significantly to the arsenal of detection methods and to the range of applications that can be dealt with successfully.

1.2 OBJECTIVES OF PHASE I

The objective of the Phase I effort was to assess the efficacy of wavelet techniques for selecting *features* upon which an adaptive classifier could base its decisions regarding abnormal changes in system behavior. For reliable, robust classification with low false alarm rates, these features must be:

- *high energy* in at least one case (normal or failed) in order to persist even in the presence of environmental noises or transient disturbances; and
- *statistically significant* in separating two or more cases from one another in order to contribute meaningful information to the pattern classifier.

Our objective was not to develop new classification techniques, but rather to modify off-the-shelf artificial neural network (ANN) (Lau and Widrow, 1990a, 1990b) technology as needed to integrate it with a front-end feature extractor based on wavelet techniques.

Wavelets offer many different ways to access the structure of a signal in time/scale space. The *continuous wavelet transform* (CWT) (Ruskai, Beylkin, et al., 1992; Daubechies, 1990; Mallat, 1989a, 1989b; Meyer,1988) converts a time signal into an image, from which features can be extracted using image processing techniques. The *wavelet packet transform* (WPT) (Coifman and Wickerhauser, 1992; Coifman et al., 1990) derives coefficients of wavelet basis functions that characterize time/scale energy distribution in a much more flexible manner than *discrete Fourier transforms* (DFTs) permit. Variations on the WPT permit the selection of subsets of an overcomplete set of basis functions to find the most significant elements of a signal. More recent extensions to wavelet techniques, presented under the general classification of multiscale signal processing, create even more options for feature characterization. Our initial goal was to select several alternatives, and to compare their performance and computational requirements in the context of whatever data were available. Because of time and budget constraints, however, we limited our comparison to CWTs and WPTs.

Because the dominant challenge in failure detection problems is to identify a concise yet distinctive set of features on which the detection/classification process can be made to depend, our emphasis in Phase I was on the back-end of the feature-selection process, i.e., we assumed that the full set of wavelet transform coefficients was already available, and then determined which were most critical to good performance of an ANN classifier. Helicopter gearbox and shipboard pump accelerometer data, supplied by the Navy, were passed through a CWT or WPT preprocessor, and then used to train ANN classifiers. Statistics on false alarm rates, miss detections, and misclassification errors were used to quantify the performance of the proposed methodology.

1.3 OVERVIEW OF PHASE I RESULTS

Phase I of this effort clearly demonstrated the feasibility of incipient fault detection for vibrating systems not only for bench test conditions (helicopter gearbox) but also for mild operating conditions (condensate and fire pumps). Remarkable Phase I results were obtained by using a balanced combination of CWTs and ANNs. We used the CWT to select features for an ANN classifier. The wavelet transform provided enough visibility into fault signals to allow us to reduce the size of the feature set to 10-15 features. We used a low-dimensional, conventional ANN classifier (Widrow et al., 1988) with rejection of ambiguous classifications. We achieved 0.000 probability of false alarm, 0.000 probability of missed detection, and < 0.04 probability of deferral (to a subsequent feature vector) for all three data sets provided by ONR. *The major product of out Phase I work is a single methodology to identify robust features that lead to these performance levels.*

1.4 REPORT ORGANIZATION

Section 2 describes the major components of the technical approach and the main results of this Phase I effort. Section 3 presents the main conclusions and recommendations for future effort. Appendix A contains an overview of the CWT and presents some examples to give the reader insight into the time-frequency information provided by the CWT based on the Kiang wavelet used in this work.

SECTION 2

TECHNICAL APPROACH TO PHASE I

This section describes the data used, the major components of the technical approach, and the main results obtained for the three types of vibrating systems considered in this work. The most important steps of the technical approach are illustrated with selected examples based on the available data sets and using illuminating color plates (located just before the body of this report).

2.1 DATA USED IN PHASE I

For this research, ONR supplied data (through the Naval Command, Control, and Ocean Surveillance Center, NCCOSC) for three vibrating systems: helicopter gearbox, condensate pumps, and fire pumps. These data are from accelerometers that measure vibrations at one or more places on the case of the vibrating mechanism.

The gearbox data (from a TH-1L helicopter intermediate (42-degree), relatively simple, gearbox) consisted of vibration readings (sampled at 48 kHz) from two accelerometers (channels 5 and 6, oriented with and orthogonal to the bearing load zones, respectively) mounted on the gearbox output end for six separate fault conditions: no defect (ND), bearing inner race fault (IR), bearing rolling element fault (RE), bearing outer race fault (OR), gear spall fault (SP), and gear 1/2 tooth cut fault (TC). This is a subset of the "Hollins data base," developed by Mark Hollins of the Naval Air Test Center (NATC). The pump data consisted of vibration readings (sampled at 50 kHz) from two accelerometer triads (axial, radial, tangential) mounted on the motor and pump ends of the assembly, one triad on each end. The condensate pump data consisted of eight data segments that included two fault types and unfailed data. The fire pump data consisted of 16 data segments that included four fault types and unfailed data. Unlike the helicopter data, which are bench test data with seeded faults, the pump data were obtained from shipboard pumps operating under relatively mild conditions.

Tables 2-1 and 2-2 present all the information we were given on the condensate and fire pump data, respectively. In both cases, fault code 0 is used to label good pumps (normal cases). Also, fire pump fault codes 3A and 3B denote the same fault type on different pumps.

TABLE 2-1. CONDENSATE PUMP INFORMATION

Segment Number	Pump ID	Fault Code	RPM
1	CP-1A	1	900
2	CP-1B	0	900
3	CP-4A	2	885
4	CP-4B	0	885
5	CP-2A	0	890
6	CP-2B	0	890
7	CP-3A	0	892
8	CP-3B	2	892

TABLE 2-2. FIRE PUMP INFORMATION

Segment Number	Pump ID	Fault Code	RPM
1	FP-9	3A	3576
2	FP-3	0	3585
3	FP-2	0	3585
4	FP-1	0	3585
5	FP-4	0	3575
6	FP-5	0	3580
7	FP-6	0	3580
8	FP-5A	0	3585
9	FP-6A	4	3580
10	FP-13A	0	3590
11	FP-12	5	3585
12	FP-13	3B	3580
13	FP-14	6	3570
14	FP-17	0	3585
15	FP-16	0	3585
16	FP-15	0	3588

For our test systems we used only one channel, from one sensor—we deferred fusion of results from multichannel data to Phase II. For the gearbox system, only channel 5 was used for all conditions. For both pump systems, only the axial component of the pump-end accelerometer triad was used. In a sense, we deliberately made the Phase I problem harder by ignoring some sources of information in order to demonstrate the power of wavelet techniques, or lack thereof, on a fault detection/classification problem more difficult than one would expect to encounter in the field under more severe conditions.

2.2 SYSTEM STRUCTURE

We adopted the conventional architecture of an adaptive classifier (Fig. 2-1): a real-time preprocessor to focus the information about the state of a system into a low-dimensional feature vector, followed by an adaptive pattern analyzer to map feature vectors into detections and classifications. Our work emphasized the development of the preprocessor, using the insight offered by recent advances in the mathematics of wavelets.

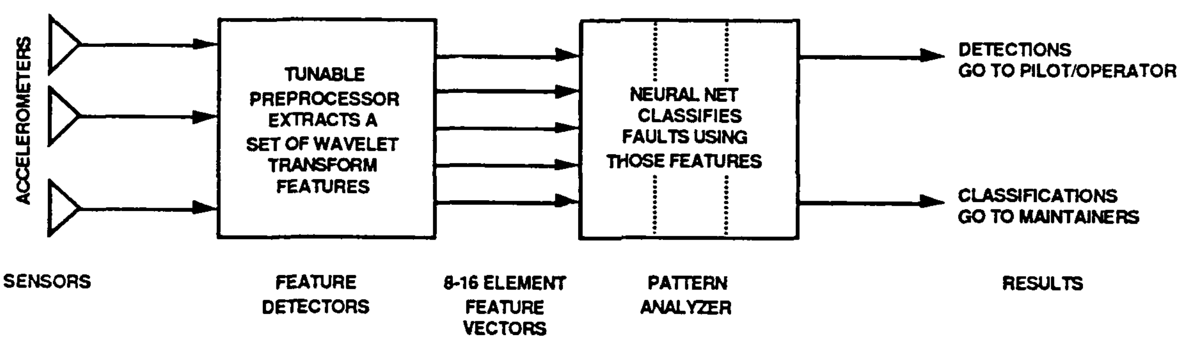

Figure 2-1. Incipient Fault Detection and Classification System Structure

Our Phase I proposal indicated that our approach of choice was to use the WPT (Coifman and Wickerhauser, 1992; Coifman, Meyer, et al.; 1990) to identify locations in time/scale space indicative of faults. This approach met with less than expected success. With hindsight, coupled with analysis of considerable performance data generated during the project, we believe that the WPT is most appropriate for problems where classification depends on the timing and internal structure of *transient* events, such as classifying biological sounds in the sea. They offer considerably less advantage in fault analysis of vibrating systems, where the signatures are of relatively long duration and statistically quite stationary. However, we reserve interest in the WPT for detecting incipient faults in systems that emit transient signals as part of their normal operation (e.g., heavy duty mechanical or electrical switching systems).

As an alternative, we turned to the CWT. Like other image-visualization techniques such as the sonogram, lofargram, or waterfall display, the CWT converts a one-dimensional signal into a two-dimensional image, using substantial computational resources. Sub-bands of the CWT can be evaluated quite efficiently in a preprocessor, however. Therefore, we use the full-blown CWT during the *design process* to identify a few bands that differentiate among cases, and only implement actual feature detectors for those specific bands—*to focus the information available in a signal into a small set of features*. This allows us to find very small feature vectors (of the order of 10 – 20 elements) that nonetheless yield outstanding detection and classification performance. A brief explanation of the CWT and examples of the CWTs of elementary signals are presented in Appendix A.

For the pattern analyzer, we used conventional three-level, feedforward ANNs. As will be seen, we succeeded in finding feature sets that are nearly convex and linearly separable, so we did not need complex network topologies or exotic training algorithms. (In fact, we were able to set the number of elements in the hidden layer equal to the number of output elements).

To improve performance we suppressed classification results entirely if the maximum output value was less than some multiple of the next larger output value, *deferring* the classification to the next available feature vector. We could trade deferral rate for false alarm/missed detection performance by changing this multiple. A multiple of 2.0 was adequate to eliminate all false alarms and missed detections, and kept the deferral rate below 4%.

The Phase I feature-selection process used a number of tools to develop feature sets. KHOROS signal processing routines (KHOROS Group, 1992), on a SUN computer, supported editing and preliminary analysis of raw data files. A custom Macintosh Pascal package computed the CWT, and another one extracted the selected feature vectors from the raw data. Excel, a commercial package from Microsoft, supported the statistical cluster analysis. Macintosh NeuralWorks, a commercial package from NeuralWare, was used to carry out the ANN training and testing. MATLAB, a commercial package from The Math Works, Inc., was used to compute performance metrics.

2.3 WAVELET-BASED TUNABLE PREPROCESSOR

Figure 2-2 presents the wavelet-based tunable feature extractor developed in this Phase I effort. The CWT is computed using the Kiang wavelet, which allows one to select appropriate frequency and time resolutions to extract from the CWT the features of interest. According to the signal characteristics one may have to smooth and decimate the CWT before extracting the frequency bands of interest. These bands are then parameterized to achieve better feature separability. Because of the properties of this wavelet-based feature extractor, it is not necessary to compute the entire CWT to extract a few features; only the frequency bands associated with the features of interest need to be computed. This leads to a significant reduction of computational effort if the number of selected features is relatively small (say, 10 to 20). (Note that while this preprocessor is currently implemented in software, it is a good candidate for hardware implementation—on a Macintosh Quadra 900 computer, with no optimization, it runs about 100 times slower than real time.)

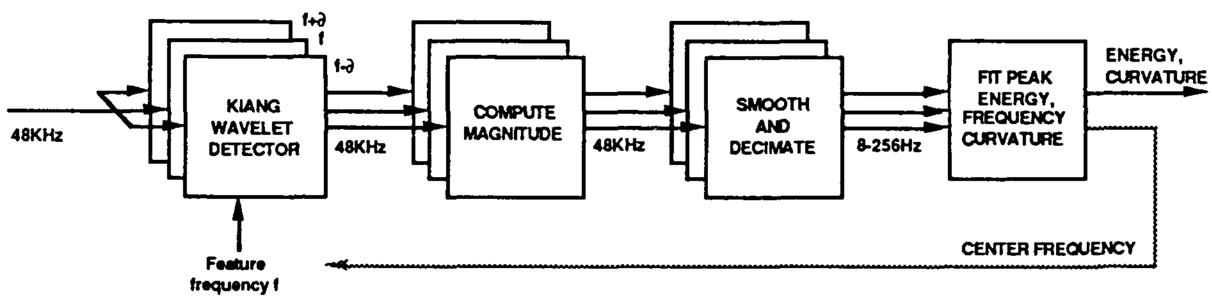

Figure 2-2. Tunable Feature Extractor

As our emphasis in Phase I was on the selection of a concise, focused feature set, we employed the visibility into time/scale space offered by the CWT. Our objective was to focus all of the potentially available time/frequency information into a small feature space, since a small, separable feature set reduces the complexity required of a classifier, and hence the risk of slow or non-convergence.

Another advantage of the CWT is its ability to isolate robust high-energy features that can be readily detected and can be suppressed only with a great deal of external energy. Our approach to feature selection was explicitly to ignore regions of time/scale space with consistently low energy, on the grounds that whatever classification might be possible using such features would not be robust to disturbances. (However, we did include one low-energy band specifically as a *disturbance detector*, where classifications would be suppressed whenever substantial energy appeared in this band.) We knew that bench test or mild-operation data is invariably cleaner than field data (it may not represent a complete range of normal operations or disturbances, and may contain artifacts not present in real data) and therefore we sought features that would serve as well in noisy environments as they do on these data sets. *Thus we are prepared to handle much more challenging data, since we address robustness at the very beginning of the feature-selection process.*

2.3.1 The Continuous Wavelet Transform

The CWT appears as an image. It is qualitatively similar to other imaging representations of a signal, such as sonograms, lofargrams, or waterfall displays. The biggest distinction is that one raster line appears in the image *for every sample in the signal*—there is no windowing of the data, and hence no artifacts in the image due to windowing. Minor distinctions include the fact that the scale axis is logarithmic, since wavelet theory involves continuous scaling of a single basis function. If one thinks of the discrete Fourier transform as the output of a bank of constant-bandwidth filters, then one can think of the CWT as the output of a bank of constant-Q (ratio of bandwidth to center frequency) filters. (The WPT, on the other hand, is like a variable-bandwidth, variable-Q filter bank (Vetterli and Herley, 1990).)

Visual analysis of the CWT reveals areas of interest in vibrational data. The raw CWT magnitudes include high-frequency artifacts that can be removed by smoothing over time. The helicopter data segments show high-energy, narrowband features overlaid by short disturbances.

The pump data segments show lower-energy, broadband features overlaid by regular impulsive disturbances. In all cases, all features are stable over time.

The CWTs presented in the color plates referenced in the sequel (and located just before the body of this document) illustrate these general observations. These CWTs use the Kiang wavelet (fine frequency, coarse time resolution). Hue encodes the log magnitude of the CWT (blue = low, red = high). Phase information is ignored. Time divisions are 62.5 msec wide. The frequency scale is logarithmic, and frequency divisions correspond to octaves.

2.3.2 Smoothing the CWT

The two images in Plate A are from the first 500 msec of channel 5 of the normal helicopter gearbox data (sampled at 48 kHz). The left image is the CWT sampled every 2 msec—the full CWT for this data segment would be 24,000 pixels high. The bright red line just below 2,048 Hz is the gear mesh fundamental. Harmonics of this fundamental appear at higher frequencies (finer scales), although there appears to be little energy in the fourth, sixth, and seventh harmonics under normal conditions. The data appears to be high-pass filtered with a cut-off frequency around 1 kHz, although some frequency lines are clearly visible below this point.

Because faults in vibrating systems impact a sensor on every cycle of the mechanism, we seek features that persist over relatively long periods of time (many cycles). While the texture of the raw CWT between harmonic lines is interesting, it would be imprudent to attempt to classify faults based on the structure of this texture. Thus for this data set, and for all other CWT images we constructed, the image was smoothed in the time dimension to suppress high-frequency textures, and enhance the stationary elements of the signal. The right image in Plate A shows the results of this smoothing. Among other things, the smoothing enhances the appearance of a secondary line just above the mesh fundamental. Also, some 5 Hz modulation on the fifth harmonic becomes more apparent.

Note that we do not present this entire image to the ANN for classification. Our goal is to identify features in the image that can be parameterized, and to compute much more concise parameter vectors from the image to submit to the net.

Note also that the CWT visualization makes the feature-selection process quite efficient. *Having developed a methodology on the gearbox data, we were able to construct low-dimensional feature sets for the two pump data sets in a matter of hours.*

2.3.3 Changing the Wavelet Basis

The two images in Plate B are from the first 500 msec of failed condensate pump data (CP-1A from Table 2-1). The left image is the analog of the smoothed helicopter CWT. It immediately shows the lack of stable, narrowband elements in the signal. Again, it would be imprudent to attempt classification based on the microstructure of the transient narrowband features in this image. Instead, we seek more broadband features.

One appeal of the CWT is that it allows one to continuously vary time/scale resolution. By changing the wavelet on which the transform is based, one can sacrifice resolution in scale—which is exactly what is necessary to find broadband features. The right image is the same data, but using a wavelet transform with 1/10th the frequency resolution (along with additional smoothing over time to suppress transients and textures). This image clearly shows the locations of high energy content—and these regions are surprisingly stable compared to those of the left image.

2.3.4 Channel Selection

The two images in Plate C are from the first 500 msec of unfailed fire pump data (FP-3 from Table 2-2). The left image is from the radial channel, the right from the axial channel. Lines for the first few harmonics of the shaft rotation frequency (about 60 Hz) are clearly visible, along with a faint harmonic series based at about 400 Hz, and some broadband signal from 300 Hz to 1,500 Hz.

As mentioned earlier, we limited Phase I processing to a single channel of data for each equipment type (so that we did not exhaust all of the potential processing gain on these clean data, and thus can offer ways to counter the additional complexity one would expect in field data). We selected the axial channel alone for further processing for the pump cases, and channel 5 alone for the gearbox case.

2.3.5 Gearbox Fault Signatures

Plate D shows segments of the CWT for the six cases of gearbox data. Each image represents 250 msec of signal, from 512 Hz to 16 kHz. Note the difference in structure of the CWT around the third harmonic of the mesh frequency (1,935 Hz)—both in breadth and texture.

2.3.6 Gearbox Fault Masks

Plate E shows the same segments of the CWT for the six cases of gearbox data, with low-energy regions masked out. Our rationale for this is to prevent using features that are weak, as they are easily compromised by disturbances or interference and hence do not contribute to high reliability detection. The technique used to mask these regions is basically local noise floor estimation across scale, masking areas that fall below that estimated floor. Note how this technique highlights the significant differences in structure around the third harmonic, and also of the 1,050 Hz line. Techniques such as this morphological filter provide quantitative evaluation of feature set performance before time and energy is spent training and testing the adaptive classifier.

2.3.7 Condensate Pump Signatures

Plate F shows the CWTs for 125 msec of the axial channel for each of the eight condensate pump data segments in Table 2-1. The log frequency scale, between 16 Hz and 16 kHz, is divided into octaves. Segment 1 contains a fault of type 1, segments 3 and 8 each contain a fault of type 2, and the rest are good pumps (normal case). Note that the CWTs show clear differences between pumps in pairs of segments 1 and 2, 3 and 4, 7 and 8. Each of these pairs includes a normal pump and a defective pump. On the other hand, segments 5 and 6, both from good pumps, display similar high-energy features.

In contrast with the helicopter gearbox data, the condensate pump CWTs contain wider high-energy regions, but they are—as in the gearbox case—relatively stable over time.

2.4 FEATURE SEPARATION

Detection and classification become exceptionally easy if the clusters of features corresponding to different cases exhibit two properties: convexity and separability. In these cases, classification becomes a matter of estimating boundaries to separate the clusters—and we can allocate one element of the hidden layer of an ANN to each cluster.

One never knows ahead of time whether or not a feature set will be convex and separable. We selected 500 msec of data from each case provided as a basis for statistical analyses of separability. We used features that essentially correspond to a few frequency slices through the CWT—energies in particular frequency bands. We selected the set of bands to use on the basis of overall energy content—recall that robust classification is possible only from features with high energy differences between cases. In the cases of the gearbox and fire pump data, with strong, clear, narrow fundamentals, we adapted the frequencies to the center of that line (using the features themselves instead of referring to external synchronization signals). For the condensate pump data (which lack such a stable reference feature and whose energies are more dissipated across frequency), we left the frequency bands constant.

After collecting candidate features, but before training an ANN, we evaluated feature cluster separations. Table 2-3 shows the maximum separations between all pairs of clusters in terms of Fisher coefficients—essentially distances normalized to units of standard deviations. This

TABLE 2-3. PHASE I CLUSTER SEPARATIONS, HELICOPTER DATA

	ND	IR	RE	OR	SP	TC
Normal	0.00	5.67	10.79	9.79	11.14	11.28
Inner Race	5.67	0.00	4.13	5.25	13.31	6.05
Rolling Element	10.79	4.13	0.00	2.07	5.89	3.45
Outer Race	9.79	5.25	2.07	0.00	7.18	3.29
Gear Spall	11.14	13.31	5.89	7.18	0.00	8.61
Tooth Cut	11.28	6.05	3.45	3.29	8.61	0.00

table was obtained by computing for each feature vector the Fisher coefficients between all fault condition pairs, and then selecting the maximum coefficient across all feature vectors for every fault condition pair. Any pair of clusters more than three or so units apart should be readily separable by an ANN classifier.

We found that the CWT features provide good separation between cases. Below is a summary of the main characteristics of these features for each of the test systems.

Helicopter gearbox feature vectors (channel 5) contain heights of narrowband (1/35) octave lines

- center frequencies adapt to changes in fundamental mesh frequency
- lines were selected at first six harmonics of mesh frequency, plus two other frequencies suggested by the CWT (0.5525*f, 2.7*f , where f is the fundamental mesh frequency)
- feature clusters are nearly ellipsoidal
- minimum feature cluster separation is 2.07 standard deviations (outer race/rolling element)

Condensate pump feature vectors (axial channel) contain energies in wider regions (1/6 octave)

- center frequencies are fixed over time (and cases)
- bands were selected at octave intervals (32-1,024 Hz), and at 1/4 octave intervals within high energy octaves (64-128 Hz, 512-1,024 Hz)
- feature clusters are nearly ellipsoidal
- minimum feature cluster separation is 5.87 standard deviations (CP-1A/CP-4B)

Fire pump feature vectors (axial channel) also contain both narrowband and broadband features

- center frequencies adapt to changes in fundamental shaft frequency (limited to within 2% nominal)
- narrow bands were selected at first 8 shaft harmonics, and at octave intervals within high energy octaves (512-2,048 Hz)
- some feature clusters show some suspiciously high correlation (probably due to clipping?)
- minimum feature cluster separation is 2.23 standard deviations (Fault 3/Fault 6)

The following three subsections graphically illustrate with color scattergrams the striking feature separation for some feature pairs and all fault conditions for each of the test systems examined in this work.

2.4.1 Gearbox Separation

Plate G presents the feature clusters computed from 3 seconds of gearbox data across all six cases. Feature vectors can be obtained every 10 msec, so there are about 300 sample vectors here. This plate shows the projection of the feature clusters onto a two-dimensional subspace defined by the power found in the second harmonic of the mesh frequency, and at a subharmonic line around 1,050 Hz. Note that: 1) all of the feature clusters appear convex, 2) the normal case is well separated from the fault cases by these two features alone, and 3) several pairs of faults can be separated as well. Other pairs of features provide different kinds of separation, but all show convex clusters.

2.4.2 Condensate Pump Separation

Plate H presents the feature clusters from 0.5 second of condensate pump data across all eight cases. There are about 30 feature vectors per case. It shows the projection of the feature clusters onto the two-dimensional subspace defined by the power near 615 Hz, and near 1,024 Hz. Again, note that: 1) all of the feature clusters appear convex, 2) the normal cases are well separated from the fault cases by these two features alone, despite being more diffuse due to variations among units, and 3) type 1 and type 2 faults can be clearly separated as well. Other pairs of features provide different kinds of separation, but all show convex clusters.

2.4.3 Fire Pump Separation

Plate I presents the feature clusters computed from 0.5 second of fire pump data across all 16 cases. There are about 30 feature vectors per case. It shows the projection of the feature clusters onto the two-dimensional subspace defined by a narrow band around the seventh harmonic of the shaft rate and a wider band around 2,048 Hz. Note some suspicious characteristics of these clusters. The seventh harmonic of the normal data seems to be limited by a floor at 42 dB below the shaft fundamental, making detection and classification of fault type 3A (light blue squares—FP-9 from Table 2-2) exceptionally easy. Also, for three of the test cases (one normal and two fault), the values of these features are exactly 6 dB apart, a relationship that is highly unlikely in truly random data. *It is important to supplement the power of data-driven approaches with some understanding of the physics of the system under study*—why classification regions are the way they are—*in order to gain confidence that the classification logic is truly robust to any artifacts that may be in the training data.*

2.4.4 Artifacts in Training Data

Through an example, this subsection illustrates the need to select features for classification based not only on their separation but also on the physics of the fault mechanism. The risk is that data presented to the ANN classifier may contain variations upon which classification may be based, but which bear no causal relationship to fault mechanisms. For example, we selected a feature at the sixth harmonic of the mesh frequency for the gearbox data to serve as a disturbance detector. In this frequency range (12 kHz), there is very little energy in any of the data unless a disturbance is present. Our idea was that if a feature vector with relatively high energy (> 45 dB below the power in the mesh fundamental) were presented to the neural net, it would result in an ambiguous classification and any output deferred until the disturbance subsides.

Quite another thing happened. Plate J shows the gearbox feature clusters projected onto the subspace defined by the fifth and sixth harmonic power levels. Note the obvious separation between {Normal, Inner Race}, {Outer Race, Rolling Element, Tooth Cut}, and {Gear Spall}. The fact that the *magnitudes* of sixth harmonic power levels are so small suggests that this frequency is near a zero in the transfer function between the vibrating mechanics and the sensor. The fact that their *variation* is so small suggests that the energy in this band is largely background energy, or conveyed through a convoluted transmission path. In either case, *the variations among cases are unlikely to be caused by the faults themselves, but rather by the process of inserting and removing faults*. An adaptive classifier will happily use the power level at the sixth harmonic to separate the Normal case from, say, the Gear Spall. Only additional insight into the physics of the transmission mechanism, or a set of data including several insertions of the same fault, would prevent field deployment of a classifier that treats this artifact as a valid source of information.

2.4.5 Guidelines for Finding Robust Feature Sets

Based on this Phase I effort we have developed a set of guidelines for finding robust feature sets in CWT images. These guidelines can be summarized as follows:

- Be sure that features are robust to external disturbances: we seek high energy content features to be derived from morphological filtering on the scale axis of the CWT, with narrow bandwidths to reduce their sensitivity to impulsive disturbances. Features must also be redundant to exploit the correlation among features, and they must be frequently computed as permitted by the largest time constant in the preprocessor.

- Be sure that features are diverse: it is desired to include features across a wide range of frequencies, for example, the first six harmonics of important narrowband vibrations (gearbox) or octave samples of broadband components (pumps). In addition, it is desired to include one or more low-energy features to support disturbance rejection.
- Be sure that features distinguish normal from abnormal conditions: for this we need to compute statistics (mean, standard deviation) on each CWT bin and look for significant differences that will lead to features with high discriminating power.

2.5 ARTIFICIAL NEURAL NETWORK CLASSIFIER

For an adaptive classifier we used a feedforward ANN of the back-propagation type with one input layer, one hidden layer, and one output layer. The number of processing elements (PEs) in the input layer varied with the vibrating system between 12 and 15. The number of output PEs also varied with the vibrating system, according to the number of fault conditions, including the normal cases. Because of the convexity and linear separation of the feature vector clusters for all the vibrating systems of Phase I, the number of hidden layer PEs was set equal to the number of output PEs. Table 2-4 presents the number of PEs per layer for each of these systems.

TABLE 2-4. NUMBER OF PROCESSING ELEMENTS PER ANN LAYER

LAYER	HELICOPTER GEARBOX	CONDENSATE PUMPS	FIRE PUMPS
Input	15	12	13
Hidden	6	8	16
Output	6	8	16
Total PEs	27	28	45

For the design, training, and testing of the ANNs we used a commercial software package, NeuralWorks (NeuralWare, 1992), running on a Macintosh platform. For each of the three systems, convergence to the specified RMS error of the difference between the desired and the actual outputs occurred relatively fast—after between 5,000 and 10,000 random presentations of the feature vectors included in the training set. *Given the simplicity of the ANNs used in this work, their small size, and the excellent feature clusters separation made possible by the judicious utilization of the CWT, no sophisticated training algorithms were required.*

From the test set results, we computed the following measures of effectiveness for each vibrating system: probability of false alarm, probability of missed detections, probability of misclassification, and probability of deferral. *Probability of false alarm* is the probability that a fault is announced when there is no fault present. *Probability of a miss detection* is the probability that no fault is announced when there is a fault present. *Probability of misclassification* is the probability that a fault type is announced when a different fault type is present. *Probability of deferral* is the probability that the classifier defers a decision when a case for decision (a feature vector) is presented to it.

For the purpose of this work, a feature vector leads to an ambiguous situation when the absolute difference between the two largest competing outputs is less than some specified tolerance. In these cases, the classifier refuses to announce a decision and considers the next feature vector. The consequence of this deferral is to decrease the probabilities of false alarm and missed detections, and to increase the time delay for a classifier decision. For instance, feature vectors for the gearbox system are computed every 10 msec, so the price paid in time delay for each deferral (or rejection) is only 10 msec.

The computation rate of feature vectors depend on the time constants selected for the wavelet preprocessor. Extracting feature vectors at a period exceeding the largest of these time constants leads to statistical independence. This period can be quite short. Table 2-5 presents the maximum time constants and number of feature vectors per second allowed by such time constants for the three vibrating systems of Phase I.

We deliberately avoided using all available techniques to solve the Phase I problem. We arbitrarily limited ourselves to single-channel processing, with instantaneous classification, in order to have some additional processing techniques available to deal with expected additional complications contained in real data.

TABLE 2-5. PREPROCESSOR TIME CONSTANTS AND FEATURE VECTOR RATES

UNIT	MAX TIME CONSTANT	FEATURE VECTORS PER SECOND
Helicopter gearbox	10 msec	100
Condensate pump	25 msec	40
Fire pump	25 msec	40

One of these techniques is temporal fusion. We compute feature vectors as fast as possible while maintaining statistical independence (i.e., at a rate limited by the longest time constant in the preprocessor). For the gearbox data, we computed feature vectors every 10 msec, and classified each and every one. To robustify the classification process against transient disturbances, we can simply compare output classifications over a window of, say, 100 feature samples (one second of data). If, say, fewer than 95 of the classifications agree, we assume a disturbance (and not a fault) is present. This dramatically reduces the theoretical probability of false alarm, at the cost of an additional second of delay in producing a warning—a very attractive tradeoff for most situations.

2.6 FAULT DETECTION AND IDENTIFICATION RESULTS FROM PHASE I

Given the preceding insight into the derivation of high-energy wavelet features and the convex, separable clusters they form in feature space, it should be no surprise that good classification results are possible. *Providing a feature set that captures the important discriminants between normal operation and faults vastly simplifies the problem of designing an adaptive classifier that achieves good performance.*

The performance results for each of the test systems are presented in Table 2-6. The acceptance threshold is the ratio between the maximum output value and the next larger output value for a given feature vector. The complexity value is the total number of PEs in the corresponding ANN. For the gearbox and the condensate pumps systems, the test set was independent from the training set; for the fire pumps data these two sets were the same.

TABLE 2-6. PHASE I PERFORMANCE RESULTS

	GEARBOX	CONDENSATE PUMP	FIRE PUMP
TRAINING SET SIZE	1125	240	480
TEST SET SIZE	6750	1400	4800
ACCEPTANCE THRESHOLD	1.4	1.2	2.0
PROBABILITY OF FALSE ALARM	0.000	0.000	0.000
PROBABILITY OF MISSED DETECTION	0.000	0.000	0.000
PROBABILITY OF DEFERRAL	0.035	0.020	0.020
PROBABILITY OF MISCLASSIFICATION	0.046	0.000	0.000
COMPLEXITY	27 PEs	28 PEs	45 PEs

The performance results in Table 2-6 clearly show that *the wavelet feature sets selected above permit perfect detection performance with low deferral rates.* While these results are pleasing, we feel that an even more important principle has been demonstrated. *We used exactly the same method to find features for the pumps as we used for the gearbox data.* There was no trial and error for the pump classifiers—these results are from the very first feature sets we picked. This offers limited but important evidence that our results are not accidental—that *we have a methodology to analyze data from vibrating systems and derive small, focused feature sets that support high-confidence fault detection and classification.*

SECTION 3

CONCLUSIONS AND RECOMMENDATIONS

3.1 CONCLUSIONS FROM THE PHASE I EFFORT

Our Phase I performance results speak for themselves. *Incipient fault detection and classification on bench-test and mild-operation data is quite feasible*, even without exploiting many additional techniques available. The CWT provides images from which feature selection is easy. These features are simple and robust (high energy, narrow bandwidth).

There are no technological impediments to practical implementation. The selected wavelet features can be computed using off-the-shelf digital filtering hardware. ANNs can be trained and employed using off-the-shelf techniques.

The helicopter gearbox data used was bench test data. Bench test data cannot support a complete characterization of normal operating regimes and cannot include a complete set of disturbances to be encountered in the field. Moreover, seeded fault data cannot span the complete set of possible failures, and may contain artifacts that assist detection. In summary, *bench tests are not reality*. Reality is much less controlled, and hence much less predictable. Although the pump data were obtained under actual operating conditions, these conditions were relatively benign and do not include all possible operating conditions. Therefore, the principal conclusion we draw from Phase I is to greet these results (or any Phase I results) with skepticism concerning their applicability to field systems, and that *any Phase II effort must focus on support for analysis and design with real data* under widely varied operating regimes.

We are well positioned to make the transition to real data in Phase II because of the methodology we established in Phase I. *We have a specific procedure for selecting low-dimensional, robust feature sets*—and have demonstrated its efficacy on the two pump data sets.

We deliberately avoided using all available processing techniques to solve the Phase I problem—we wanted to be sure that a clean problem could be solved with simple techniques. We arbitrarily limited ourselves to single-channel processing, with instantaneous classification. *This makes available some additional processing techniques to meet the challenges posed by using real data, and we propose to exploit them to the fullest in Phase II.*

3.2 PHASE II RECOMMENDATIONS

Phase II must address two classes of issues: 1) continued development of integrated wavelet/ANN techniques for incipient fault detection, and 2) preparation for transition to Phase III through the development of: a) a generic off-line software design suite and b) a hardware, real-time feature extraction capability.

While Phase I demonstrated the feasibility of achieving good fault detection and classification performance, the methods employed were neither terribly efficient nor able to incorporate all sources of information. The major technical issues to be resolved relate to these two areas—making the feature selection more efficient, and being able to fuse other sources of failure information (e.g., multichannel data) into the classifier.

To achieve efficiency, one must integrate the software tools used to develop the Phase I results (CWT routines, statistical analysis of candidate features, creation and editing feature sets, ANN training and testing, ANN performance characterization, and deferral threshold trades), and extend the technological bases of several of them. Phase II must consolidate all of the requisite functionality currently provided by multiple software packages on two hardware platforms into a single design package. KHOROS provides a perfectly acceptable shell for all of these functions. The objective of Phase II in this area is thus to migrate all of these functions, along with any new ones developed in the process of resolving the technical issues raised below, into a unified wavelet feature analysis and selection software environment.

Some of the technical areas demanding additional research attention include the following: Can one perform the statistical evaluation of wavelet features directly in wavelet-transform space (i.e., for all candidate features) rather than for selected features in a post-transform analysis? Can one apply advanced multiscale estimation techniques to extract different features at different time scales? How effective might be some additional wavelet feature types (e.g., wavelet packet coefficients for switching systems, or multiscale autoregressive model identification)? Can recent work on the interpretation of ANN outputs as likelihood functions be directly linked to the statistical performance analysis currently done in feature space?

In addition, one is not limited to single-channel accelerometer data in many potential applications. How can one process additional channels? Is there advantage to using a vector wavelet transform to process all channels simultaneously? Can any statistical techniques derived for use in wavelet transform space be adapted to vector transforms? Is there a need for different techniques on different channels? To what extent can accelerometers mounted on the supports of a vibrating system, rather than on its casing, supply information about environmental vibrations and disturbances? How should this information influence the feature-selection process?

Finally, there are several issues related to incorporating incipient failure detectors into a genuine Navy maintenance concept. What are the relative merits of detection alone vs. detection and classification? What are acceptable false alarm rates? Are there other meaningful outputs from an incipient fault detector (e.g., rate of development of an anomaly) that can be useful to an operator? And, of course, what response time is required, and how can that time be used to process a series of feature vectors in order to reduce false alarms and increase detection reliability.

Transition to Phase III demands a generic product that can select and extract wavelet features for any specific application. This product must consist of two parts. The first is a design capability, which an engineer can use to select feature sets relevant to any particular application. The second is a real-time feature extractor, which (after setting some parameters to values determined during the design effort) will produce digitally sampled feature vectors in real time.

Phase II must also address computational efficiency. The wavelet feature extractor in Phase I executed about 100 times slower than real time. Most of the functions are simple filtering operations, and many can be done in parallel (one wavelet feature per channel). Since the feature extraction relies on one generic set of computations for each feature, parameterized to suit a particular application, now is the time to migrate these functions into simple, generic hardware. Design options range from the use of standard DSP chips to a hybrid analog front-end/microprocessor backend. Since accelerometers (and most foreseeable other sensors) supply data at audio frequencies, there should be no need to push the state of the hardware art in signal processing hardware design—off-the-shelf components should easily provide the necessary performance and physical reliability.

Thus, Phase II has the central objective of producing a *generic capability to rapidly design and implement incipient failure detectors* for a wide range of Navy applications. This capability should consist of an off-line design software suite for feature selection, ANN training, and performance evaluation. It should also include a hardware element that can be tuned to extract a range of wavelet features, in real time, so that the same hardware element can be used for a variety of applications.

Having a capability to demonstrate incipient fault detection is not enough, however. There must be ample evidence that the capability delivers products that work. Therefore, a final objective of Phase II must be not only to *develop* the wavelet feature selection and extraction capability, but to *demonstrate it on large volumes of real data.*

REFERENCES

ALPHATECH, Inc., "Condition-Based Machinery Maintenance—Phase II", Technical Proposal P-3800-98, ALPHATECH, Inc., Burlington, MA, 25 November 1992.

Basseville, M., "Detecting Changes in Signals and Systems—A Survey," *Automatica*, Vol. 24, 1987, pp. 309-326.

Coifman, R.R. and M.V. Wickerhauser, "Entropy-Based Algorithms for Best Basis Selection," *IEEE Trans. Inform. Theory*, Vol. 38, No. 2, March 1992, pp. 713-718.

Coifman, R.R., Y. Meyer, S. Quake, and M.V. Wickerhauser, "Signal Processing and Compression with Wave Packets," preprint, April 1990.

Daubechies, I., "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Trans. Inform. Theory*, Vol. 36, 1990, pp. 961-1005.

KHOROS Group, "Khoros Software," Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, 1992.

Lau, C. and B. Widrow, Special Issue on Neural Networks, I: Theory and Modeling, *Proc. IEEE*, September 1990.

Lau, C. and B. Widrow, Special Issue on Neural Networks, II: Analysis, Techniques, and Applications, *Proc. IEEE*, October 1990.

Mallat, S.G., "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 11, July 1989, pp. 674-693.

Mallat, S.G., "Multifrequency Channel Decompositions of Images and Wavelet Models," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 37, December 1989, pp. 2091-2110.

Meyer, Y., *Ondelettes et Operateurs*, Hermann, 1988.

NeuralWare, Inc., "NeuralWorks Professional II/PLUS Software", Pittsburgh, PA, 1992.

Ruskai, M.B., G. Beylkin, et al., (Eds), *Wavelets and Their Applications*, Jones and Bartlett Publishers, Boston, 1992.

Vetterli, M. and C. Herley, "Wavelets and Filter Banks: Relationships and New Results," *Proc. ICASSP*, Albuquerque, NM, 1990.

Weiss, J.L., and J.Y. Hsu, "Integrated Restructurable Flight Control Demonstration Results," NASA CR-178305, NASA Langley Research Center, May 1987.

Widrow, B., R. Winter, and R. Baxter, "Layered Neural Nets for Pattern Recognition," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 36, 1988, pp. 1109-1118.

Willsky, A.S., "A Survey of Design Methods for Failure Detection in Dynamic Systems," *Automatica*, November 1976, pp. 601-611; also presented as an invited paper at the *MIT, NASA-Ames Workshop on System Reliability for Future Aircraft*, Cambridge, MA, August 1975 (NASA Report NASA CP-003); also in *AGARDograph* No. 224 on Integrity in Electronic Flight Control Systems.

ADDITIONAL RELEVANT PUBLICATIONS BY PHASE I TEAM MEMBERS

Basseville, M. A. Benveniste, A.S. Willsky, and K.C. Chou, "Multiscale Statistical Processing," *Proc. Int. Symp. Math. Theory of Networks Syst.*, Amsterdam, June 1989.

Basseville, M., A. Benveniste, and A.S. Willsky, "Multiscale Autoregressive Processes, Parts I and II," *IEEE Trans. Signal Processing*, August 1992.

Basseville, M., A. Benveniste, K.C. Chou, S.A. Golden, R. Nikoukhah, and A.S. Willsky, "Modeling and Estimation of Multiresolution Stochastic Processes," *IEEE Trans. Inform. Theory*, Vol. 38, No. 2, March 1992, pp.766-784.

Chou, K.C. A.S. Willsky, A. Benveniste, and M. Basseville, "Recursive and Iterative Estimation Algorithms for Multi-Resolution Stochastic Processes," *Proc. 28th IEEE Conf. Decision Contr.*, Tampa, FL, December 1989.

Chou, K.C. and A.S. Willsky, "Multiscale Riccati Equations and a Two-Sweep Algorithm for the Optimal Fusion of Multiresolution Data," *Proc. 29th IEEE Conf. Decision Contr.*, Honolulu, HI, December 1990.

Chou, K.C., S. Golden, and A.S. Willsky, "Modeling and Estimation of Multiscale Stochastic Processes," *ICASSP-91*, Toronto, April 1981.

Chow, E.Y. and A.S. Willsky, "Analytical Redundancy and the Design of Robust Failure Detection Systems," *IEEE Trans. Automat. Contr.*, Vol. 29, July 1984, pp. 603-604.

Chow, E.Y. and A.S. Willsky, "Bayesian Design of Decision and Rules for Failure Detection," *IEEE Trans. Aerospace and Electronic Syst.*, Vol. 20, No. 6, November 1984, pp. 761-774.

Chow, E.Y., X.-C. Lou, G.C. Verghese, and A.S. Willsky, "Redundancy Relations and Robust Failure Detection," in *Detection of Abrupt Changes in Signals and Dynamical Systems*, M. Basseville and A. Benveniste (Eds.), Springer-Verlag Lecture Notes in Control and Information Sciences, No. 77, 1986.

Deckert, J.C., M.N. Desai, J.J. Deyst, and A.S. Willsky, "F8-DFBW Sensor Failure Identification Using Analytic Redundancy," *IEEE Trans. Automat. Contr.*, Vol. AC-22, No. 5, October 1977, pp. 795-803.

Dowdle, J.R., M. Athans, S. Gully, and A.S. Willsky, "Optimal Estimation of Maneuvering Targets for Missile End-Game Guidance," *Proc. Amer. Contr. Conf.*, Arlington, VA, June 1982.

Eterno, J.S., J.L. Weiss, D.P. Looze, and A.S. Willsky, "Design Issues for Fault Tolerant-Restructurable Aircraft Control," *IEEE Conf. Decision Contr.*, December 1985.

Grunberg, D.B., J.L. Weiss, and J.C. Deckert, "Generation of Optimal and Suboptimal Strategies for Multiple-Fault Isolation," *Proc. IEEE Int'l. Automatic Testing Conf.*, San Francisco, CA, 3-5 November 1987, pp. 125-131.

Grunberg, D.B., P. Kapasouris, J.L. Weiss, J.C. and Deckert, "Computer-Aided Design Techniques for Test System Development: Phase II Final Report," TR-418, ALPHATECH, Inc., Burlington, MA, January 1989.

Gustafson, D.E., A.S. Willsky, J.-Y. Wang, M.C. Lancaster, and J.H. Triebwasser, "A Statistical Approach to Rhythm Diagnosis in Cardiograms," *Proc. IEEE*, Vol. 65, No. 5, May 1977, pp. 802-804.

Gustafson, D.E., A.S. Willsky, J.-Y. Wang, M.C. Lancaster, and J.H. Triebwasser, "ECG/VCG Rhythm Diagnosis Using Statistical Signal Analysis I: Identification of Persistent Rhythms," *IEEE Trans. Biomed. Eng.*, Vol. 25, July 1978, pp. 344-353.

Gustafson, D.E., A.S. Willsky, J.-Y. Wang, M.C. Lancaster, and J.H. Triebwasser, "ECG/VCG Rhythm Diagnosis Using Statistical Signal Analysis II: Identification of Transient Rhythms," *IEEE Trans. Biomed. Eng.*, Vol. 25, July 1978, pp. 353-361.

Korn, J., S. Gully, and A.S. Willsky, "Tracking of Maneuvering Aircraft by Surface-to-Air Missiles," *Proc. Amer. Contr. Conf.*, Arlington, VA, June 1982.

Lou, X.-C., A.S. Willsky, and G.C. Verghese, "Optimally Robust Redundancy Relations for Failure Detection in Uncertain Systems," *Automatica*, Vol. 22, 1986, pp. 333-344.

Mier-Muth, A.M. and A.S. Willsky, "A Sequential Method for Spline Approximation with Variable Knots," *Int. J. of System Science*, Vol. 9, September 1978, pp. 1055-1067.

Pattipati, K.R., A.S. Willsky, J.C. Deckert, J.S. Eterno, and J.L. Weiss, "A Design Methodology for Robust Failure Detection and Isolation," *Proc. Amer. Contr. Conf.*, San Diego, CA, 6-8 June 1984, pp. 1755-1762.

Pattipati, K.R., Alexandridis, M.G., and Deckert, J.C., "Efficient Sequencing of Diagnostic Tests," *Proc. 40th Meeting of the Mechanical Failures Prevention Group*, National Bureau of Standards, Gaithersburg, Maryland, 16-18 April 1985, pp. 122-135.

Pattipati, K.R., J.C. Deckert, and M.G. Alexandridis, "Time-Efficient Sequencer of Tests (TEST)," *Proc. IEEE Int. Automatic Testing Conf.*, Uniondale, NY, 22-24 October 1985, pp. 49-62.

Pattipati, K.R., J.J. Shaw, J.C. Deckert, L.K. Beean, M.G. Alexandridis, and W.E. Lougee "CONFIDANTE: A Computer-Based Design Aid for the Optimal Synthesis, Analysis, and Operation of Maintenance Facilities," *Proc. IEEE Automatic Testing Conf.*, Washington, DC, 5-7 November 1984, pp. 390-404.

Pattipati, K.R., M.G. Alexandridis, and J.C. Deckert, "A Heuristic Search/Information Theory Approach to Near-Optimal Diagnostic Test Sequencing," *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Atlanta, GA, 14-17 October 1986, pp. 230-235.

Weiss, J.L, A.S. Willsky, D.P. Looze, J.T. Crawford, and R.R. Huber, "Detection and Isolation of Control Surface Effectiveness Failures in High-Performance Aircraft," *Proc. National Aerospace and Electronics Conf.*, Dayton, OH, May 1985.

Weiss, J.L. A.S. Willsky, K.R. Pattipati, and J.S. Eterno, "Application of FDI Metrics to Detection and Isolation of Sensor Failures in Turbine Engines," *Proc. Amer. Contr. Conf.*, Boston, MA, June 1985, pp. 1114-1120.

Weiss, J.L., J.C. Deckert, and K.B. Kelly, "Analysis and Demonstration of Diagnostic Performance in Modern Electronic Systems," RL-TR-91-180, Rome Laboratory, Griffiss AFB, NY, August 1991.

Weiss, J.L., K.R. Pattipati, A.S. Willsky, J.S. Eterno, and J.T. Crawford, "Robust Detection/Isolation/Accommodation for Sensor Failures," ALPHATECH Technical Report TR-213, September 1984.

Weiss, J.L., K.R. Pattipati, A.S. Willsky, J.S. Eterno, and J.T. Crawford, "Robust Detection/Isolation/Accommodation for Sensor Failures," NASA CR-174797, NASA Lewis Research Center, September 1985.

Weiss, J.L., M. Doyle-Buckley, and J.C. Deckert, "A Prototype Test-Engineering Workstation," Proc. National Aerospace and Electronics Conf., Dayton, OH, June 1989, pp. 1937-1943.

Willsky, A.S. and H.L. Jones, "A Generalized Likelihood Ratio Approach to the Detection and Estimation of Jumps in Linear Systems," *IEEE Trans. Automat. Contr.*, Vol. AC-21, February 1986, pp. 108-112.

Willsky, A.S., "Detection of Abrupt Changes in Dynamic Systems," in *Detection of Abrupt Changes in Signals and Dynamical Systems*, M. Basseville and A. Benveniste (Eds.), Springer-Verlag Lec. Notes in Control and Information Sciences, No. 77, 1986.

Willsky, A.S., "Multiscale Stochastic Processes and Wavelet Transforms," *Lowell Conference on Wavelets*, Lowell, MA, June 1990, invited presentation.

Willsky, A.S., E.Y. Chow, S.B. Gershwin, C.S. Greene, P.K. Houpt, and A.L. Kurkjian, "Dynamic Model-Based Techniques for the Detection of Incidents on Freeways," *IEEE Trans. Automat. Contr.*, Vol. AC-25, No. 3, June 1980, pp. 347-360.

Willsky, A.S., J.J. Deyst, and B.S. Crawford, "Two Self-Test Methods Applied to an Inertial System Problem," *AIAA J. Spacecraft and Rockets*, Vol. 12, July 1975, pp. 434-437.

SELECTED BIBLIOGRAPHY ON FAULT DETECTION/ISOLATION AND WAVELETS

Basseville, M. and A. Benveniste, "Sequential Detection of Abrupt Changes in Spectral Characteristics of Digital Signals," *IEEE Trans. Inform. Theory*, Vol. 29, September 1983, pp. 709-723.

Basseville, M., A. Benveniste, and G. Moustakides, "Detection and Diagnosis of Abrupt Changes in Modal Characteristics of Nonstationary Digital Signals," *IEEE Trans. Inform. Theory*, Vol. 32, May 1986, pp. 412-417.

Basseville, M., A. Benveniste, G. Moustakides, and A. Rougee, "Detection and Diagnosis of Changes in Eigenstructure of Nonstationary Multivariable Systems, *Automatica*, Vol. 23, 1987, pp. 479-489

Benveniste, A. and J.-J. Fuchs, "Signal Sample Modal Identification of a Nonstationary Stochastic Process," *IEEE Trans. Automat. Contr.*, Vol. 30, January 1985, pp. 66-74.

Benveniste, A., M. Basseville, and G. Moustakides, "The Asymptotic Local Approach to Change Detection and Model Validation," *IEEE Trans. Automat. Contr.*, Vol. 32, 1987, pp. 583-592.

Beylkin, G., R. Coifman, and V. Rokhlin, "Fast Wavelet Transforms and Numerical Algorithms I, *Commun. Pure and Appl. Math.*.

Burt, P.J. and E.H. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Trans. Commun.*, Vol. 30, No. 4, April 1983, pp. 532-540.

Cheung, J. and G. Stephanopoulos, "Representation of Process Trends—Part I: A Formal Representation Framework," *Computers Chem. Eng.*, Vol. 14, No. 4/5, 1990, pp. 495-510.

Cheung, J. and G. Stephanopoulos, "Representation of Process Trends—Part II: The Problem of Scale and Qualitative Scaling," *Computers Chem. Eng.*, Vol. 14, No. 4/5, 1990, pp. 511-539.

Chou, K.C., "A Stochastic Modeling Approach to Multiscale Signal Processing," PhD Thesis, MIT, May 1991.

Daubechies, I., "Orthonormal Bases of Compactly Supported Wavelets," *Commun. Pure and Appl. Math*, Vol. 41, November 1988, pp. 909-996.

Desai, M. and D.J. Shazeer, "Acoustic Transient Analysis Using Wavelet Decomposition," *Neural Networks for Ocean Engineering Workshop*, Washington, DC, August 1991.

Doyle-Buckley, M. and J.L. Weiss, "ALPHATECH Program for Testability (APT) Computing Environment (ACE)," TR-362-2, ALPHATECH, Inc., Burlington, MA, January 1989.

Frank, P.M, "Fault Diagnosis in Dynamic System via State Estimation—A Survey," in Tzafestas et al. (Eds.), *Reliability and Related-Knowledge-Based Approaches*, Vol. 1, 1987, pp. 35-98, D. Reidel Publishing Company.

Friedlander, B. and B. Porat, "Detection of Transients by the Gabor Representation," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 37, February 1989, pp. 169-180.

Friedlander, B. and B. Porat, "Performance Analysis of Transient Detectors Based on a Class of Linear Data Transforms," *IEEE Trans. Inform. Theory*, Vol. 38, No. 2, March 1992, pp. 665-673.

Frisch, M. and H. Messer, "Detection of a Transient Signal of Unknown Scaling and Arrival Time Using the Discrete Wavelet Transform," *ICASSP-91*.

Glowinski, R., W. Lawton, M. Ravachol, E. Tennenbakum, "Wavelet Solution of Linear and Nonlinear Elliptic, Parabolic, and Hyperbolic Problems in One Space Dimension," *9th SIAM Int. Conf. Comp. Methods in Appl. Sci. and Eng.*, 1990.

Golden, S.A., "Identifying Multiscale Statistical Models Using the Wavelet Transform," SM Thesis, MIT, May 1991.

Goodwin, G.C., R.J. Evass, R.L. Leal, and R.A. Feik, "Sinusoidal Disturbance Rejection with Application to Helicopter Flight Data Estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 34, June 1986, pp. 479-484.

Isermann, R., "Process Fault Detection Based on Modeling and Estimation Methods—A Survey, *Automatica*, Vol. 20, 1984, pp. 387-404.

Kim, M. and A.H. Tewfik, " Fast Multiscale Detection in the Presence of Fractional Brownian Motions," *Proc. SPIE Conf. on Advanced Algorithms and Architecture for Signal Processing V*, San Diego, CA, July 1990.

Korn, J. and N.R. Sandell, "Tool Wear Sensing and Failure Detection via Modern Estimation Theory," ALPHATECH Technical Report TR-131, Burlington, MA, March 1982.

Korn, J. and S.W. Gully, "On-Line Wear Estimation and Failure Detection of Oil-Well Drill Bits," ALPHATECH Technical Report TR-271, Burlington, MA, November 1985.

Lawton, W.M., "Wavelet Discretization Methods for Surface Estimation and Reconstruction," *SPIE/SPSE Symp. on Elec. Imaging Sci. and Tech.*, Santa Clara, CA, February 1990.

Le Vey, G., "Analyse Modale et Surveillance Vibratoire des Machines Tournantes," PhD Thesis, University of Rennes, France, 1988.

Ljung, L. and T. Söderström, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA, 1983.

Mallat, S. and W.L. Hwang, "Singularity Detection and Processing with Wavelets," *IEEE Trans. Inform. Theory*, Vol. 38, No. 2, March 1992, pp. 617-643.

Mallat, S.M. and S. Zhang, "Compact Image Coding from Edges with Wavelets," *ICASSP-91*.

Mallat, S.M., "Zero-Crossings of a Wavelet Transform,"*IEEE Trans. Inform. Theory*, July 1981.

Moustakides, G., M. Basseville, A. Benveniste, and G. Le Vey, "Diagnosing Mechanical Changes in Vibrating Systems," IRISA (Institut de Researche en Informatique et Systèmes Aléatoires) Report 436, Rennes, France, October 1988.

Pati, Y.C. and P.S. Krishnaprasad, "Analysis and Synthesis of Feedforward Neural Networks Using Discrete Affine Wavelet Transforms," preprint, Univ. Maryland, 1991.

Pentland, A., "Finite Element and Regularization Solutions Using Wavelet Bases," MIT Media Lab Vision and Modeling Group TR-143, June 1990.

Porat, B. and B. Friedlander, "Adaptive Detection of Transient Signals," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 34, December 1986.

Rougée, A., "Détection de Changements dans les Paramètres AR d'un Processus ARMA Vectoriel: Application à la Surveillance des Vibrations," PhD Thesis, University of Rennes, 1985.

Tewfik, A.H. and M. Kim, "Correlation Structure of the Discrete Wavelet Coefficients of Fractional Brownian Motions," *IEEE Trans. Inform. Theory*, Vol. 38, No. 2, March 1992, pp. 904-909.

Tuteur, F.B., "Wavelet Transformations in Signal Detection," *ICASSP-88*.

Weiss, J.L., and J.Y. Hsu, "Design and Evaluation of a Failure Detection and Isolation Algorithm for Restructurable Control Systems," NASA CR-178213, NASA Langley Research Center, March 1987.

Weiss, J.L., and J.Y. Hsu, "Expanded Envelope Concepts for Aircraft Control-Element Failure Detection and Identification," NASA CR-181644, NASA Langley Research Center, June 1988.

Weiss, J.L., J.L. Stifel, and K.S. Govindaraj, "Flight Test Results of a Control Element Failure Detection and Isolation Algorithm," *Proc. National Aerospace and Electronics Conf.*, Dayton, OH, May 1986.

Wickerhauser, M.V., "Acoustic Signal Compression with Wave Packets," preprint.

Wornell, G.W. and A.V. Oppenheim, "Estimation of Fractal Signals from Noisy Measurements Using Wavelets," *IEEE Trans. Signal Processing*, to appear.

Wornell, G.W. and A.V. Oppenheim, "Wavelet-Based Representations for a Class of Self-Similar Signals with Application to Fractal Modulation," *IEEE Trans. Inform. Theory*, Vol. 38, No. 2, March 1992, pp. 785-800.

Wornell, G.W., "A Karhunen-Loéve-Like Expansion for 1/f Processes via Wavelets," *IEEE Trans. Inform. Theory*, Vol. 36, No. 9, July 1990, pp. 859-861.

Zhang, Q. and A. Benveniste, "Approximation by Nonlinear Wavelet Networks," *ICASSP-91*, Toronto, May 1991.

APPENDIX A

THE CONTINUOUS WAVELET TRANSFORM

This appendix presents the basics of the Continuous Wavelet Transform, which underlies our methodology for selecting features.

A.1 WAVELET BASES

Wavelets are a new approach to an old problem: building complicated functions out of simple elements. Fourier analysis builds complicated functions out of sine and cosine functions. Wavelets use functions with limited time extent, so that they can better represent time-localized aspects of a function than can a sum of infinite-duration sines and cosines.

All wavelet analyses are based on dilations, contractions, and translations of a basic *mother wavelet*. If the dilations, contractions, and translations are orthogonal to one another, then the wavelets form an orthonormal basis which is complete in many cases. One unique characteristic of wavelet analysis is that there are an infinite number of choices for a mother wavelet, allowing one to continuously vary the tradeoff between time and frequency localization. Figure A-1 shows the Haar wavelet, on the left, and the Kiang wavelet, on the right. Figure A-2 shows dilated and contracted translates of the Haar wavelet.

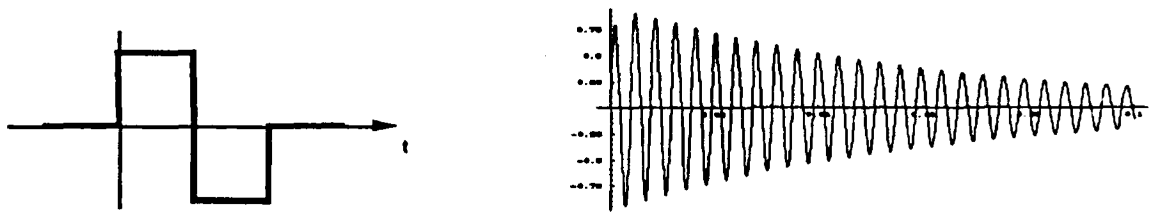

Figure A-1. The Haar and Kiang Wavelets

In its most general form, a mother wavelet is any essentially time- and band-limited function of t, subject to the uncertainty principle limitations. (The uncertainty principle states that good time localization is obtained at the expense of frequency localization, and vice versa.)

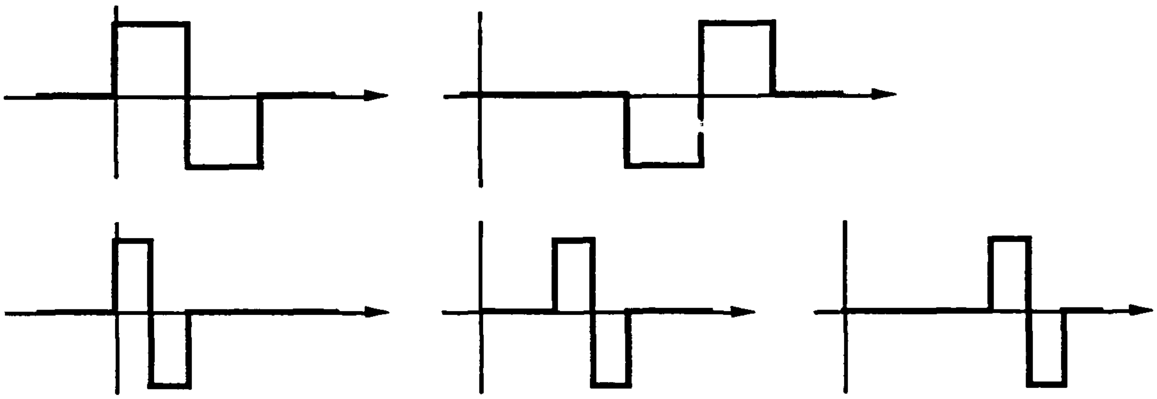

Figure A-2. Dilated Translates of the Haar Wavelet

An affine wavelet family is obtained from dilations and translations of the mother wavelet. Affine wavelets can be either continuous or discrete (usually dyadic). Dyadic wavelet families dilate by powers of two, and translate by integral multiples of power of two. Continuous wavelet families dilate by arbitrary scale factors, and translate by arbitrary amounts.

A.2 WAVELETS AND SPECTRA

Each mother wavelet has a characteristic footprint in time/frequency space—a region where it is sensitive to energy. Dilate it (or contract it) by a factor of two, and this region expands (or contracts) by a factor of two along the time axis, and translates by an octave down (or up) along the frequency axis. Translate it, and the footprint moves an equal amount along the time axis. Take a set of dilations, contractions, and translations that cover all of time/frequency space, and you have a complete basis set. Figure A-3 illustrates a typical subdivision by wavelets of the time-frequency plane.

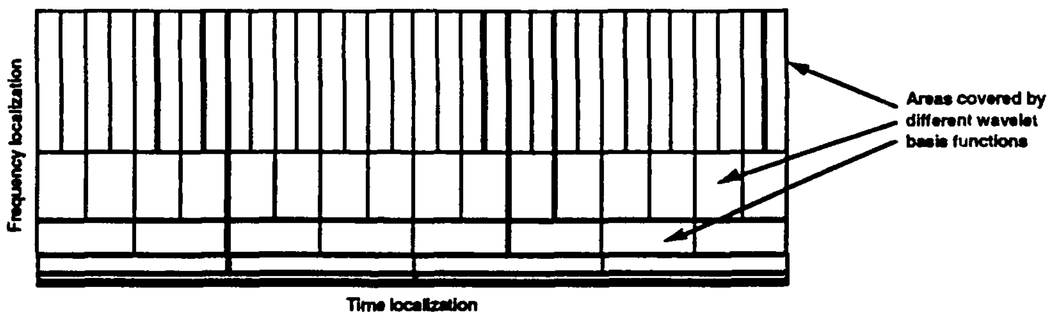

Figure A-3. Typical Subdivision by Wavelets of the Time-Frequency Plane

The exciting discoveries that kindled the recent interest in wavelets concern the existence of such complete bases that are also orthonormal, and for which the mother wavelet has finite support (is non-zero over a finite length of time).

A.3 WAVELET TRANSFORMS

Because there are an infinite number of mother wavelets, there are an infinite number of wavelet transforms. By varying the choice of the mother wavelet, one can get wavelet transforms that differ in both time and frequency localization.

An affine wavelet transform is the set of coefficients that multiply elements of a wavelet family in order to represent a time function. Affine wavelet transforms can be either continuous or discrete. Continuous wavelet transforms yield a complex-valued function of time and space (frequency). Discrete (usually dyadic) wavelet transforms yield a finite number of coefficients for an essentially bounded region of time and scale (frequency). All wavelet transforms contain implicit or explicit feature detectors (one for each basis function). Figure A-4 presents a classification of affine and non-affine wavelets based on their support (columns) and extent of their filter realizations (rows).

Weyl-Heyerdahl wavelets (non-affine)

	Bounded	Semi-infinite	Infinite
Finite	DFT Henning Hamming	Laplace (limited)	
Infinite		Bessel	Gabor Weyl

Historical interest

Affine wavelets

	Bounded	Semi-infinite	Infinite
Finite	Haar Daubechies	Laplace (limited) Kieng	
Infinite		Bessel	"Mexican Hat"

Recent results

Figure A-4. Classification of Wavelets

Wavelet selection affects preprocessor design. Efficient real-time feature extraction imposes constraints on possible mother wavelets. We require causality and a finite dimensional realization of the wavelet generator. In this effort, we constrained the choice of mother wavelet based on our need to implement a feature extractor without new leaps forward in electronics technology. In particular, we limited our choices to those for which corresponding filters have a finite dimensional realization. All of the wavelets with finite support possess this property, but must be very long in order to localize the narrowband energy so obvious in the gearbox data. Therefore, we selected a wavelet with semi-infinite support, so that it has a causal realization, and a low-dimensional implementation. This wavelet was derived from auditory nerve response data collected by Kiang in the 1960s, and we have honored him by appropriating his name for the wavelet.

A.4 SIMPLE EXAMPLES OF CONTINUOUS WAVELET TRANSFORMS

This section presents color images of Continuous Wavelet Transforms (CWTs) of some basic functions to give the reader greater insight into the time-frequency representation of signals provided by the CWT based on the Kiang wavelet. In these images, the frequency scale is logarithmic and runs vertically, increasing to the top, with frequency divisions corresponding to octaves from 16 Hz to 16 kHz. The time scale runs horizontally, increasing to the right, with time divisions 62.5 msec wide. Hue encodes the log magnitude of the CWT (blue = low, red = high). Phase information is ignored.

A.4.1 Pulse and Sine Wave

The left image in Plate K is a wavelet transform, using the Kiang wavelet, of a 1 msec pulse sampled at 48 kHz. Note the nulls at multiples of 1 kHz, reflecting the fact that the Kiang wavelet is highly oscillatory and thus almost orthogonal to pulses that extend over an integral number of complete cycles. Note that the finer scales allow very precise placement of the time that the pulse occurs, and the absence of windowing effects.

The right image is a wavelet transform of a stepped sine wave. Note that initial transients quickly give way to a very tight localization of the center frequency of the sine. Note also that the transients decay more slowly for the coarser scales at the right of the figure.

A.4.2 Superposition Examples

The left image in Plate L is a wavelet transform of the sum of the two functions in Plate K. It provides clear evidence that a single transform can localize transient events well in time, while at the same time localizing stationary frequencies.

The right image is a wavelet transform of two stepped sine waves of identical amplitude and similar frequency. It provides a dramatic visualization of beat effects, as the pattern of the peak is a periodic variation between two separate, lower-energy signals and a single, higher-energy signal. As the difference between center frequencies increases, proportionally more of each beat cycle is occupied by the area with two distinct frequency peaks. Note that the detail of the beat structure is not typically present in sonograms, lofargrams, or waterfall displays—and it is this visibility into time/frequency variations of signal structure that is the advantage of the CWT.

A.4.3 Noise Examples

The left image in Plate M is a wavelet transform of a white Poisson process emitting 1-msec pulses, with a mean interarrival time of 10 msec. Note the clear separation between pulses apparent at the finer scales at the left of the image, in contrast to the random texture at lower scales at the right. There is clearly no structure to this signal that is localized in frequency.

The right image is a wavelet transform of a white Gaussian process. The apparent concentration of energy at the finer scales (higher frequencies) is due to the fact that the bandwidths of wavelets inevitably increase as they are contracted. Thus a process with equal energy at each frequency in a Fourier transform has exponentially increasing power as frequency increases in a wavelet transform.